DOI: 10.1111/nous.12453

ARTICLE

NOÛS

The normativity of gender

Rach Cosker-Rowland

School of Philosophy, Religion, and History of Science, University of Leeds, United Kingdom

Correspondence
Rach Cosker-Rowland, School of Philosophy, Religion, and History of Science, University of Leeds, United Kingdom
Email: r.rowland@leeds.ac.uk

Abstract

There are important similarities between moral thought and talk and thought and talk about gender: disagreements about gender, like disagreements about morality, seem to be intractable and to outstrip descriptive agreement; and it seems coherent to reject any definition of what it is to be a woman in terms of particular social, biological, or other descriptive features, just as it seems coherent to reject any definition of what it is to be good or right in terms of any set of descriptive properties. These similarities give us reason to investigate the idea that, like moral thought and talk, gender thought and talk is inherently normative. This paper proposes a normative account of gender thought and talk in terms of fitting treatment. On this fitting treatment account, to judge that A is gender G is just to judge that it is fitting to treat A as a G. This account is a descriptive or hermeneutical account of our gender thought and talk rather than an ameliorative account of our gender concepts or a metaphysical account of gender properties in social metaphysics. This paper argues that other descriptive accounts of gender thought and talk face problems that the fitting treatment account overcomes.

1 | INTRODUCTION

There are some notable similarities between moral thought and talk and gender thought and talk. First, like moral disagreements, disagreements about gender seem to be deep and intractable and to outstrip descriptive agreement. Act-utilitarians and deontologists can agree about all the descriptive features of the footbridge trolley case but still disagree about whether it is wrong to push the heavy man to his death to save five others; Nozickian libertarians and Rawlsian egalitarians can agree about all the descriptive features of a society's distribution of resources, yet still disagree about whether it is a just distribution. Similarly, gender-critical feminists and trans-inclusive feminists can agree about all the biological and social features of trans women yet still disagree about whether trans women are women. Second, just as it seems coherent to doubt any definition of 'right' and 'good' in terms of natural or social features, similarly it seems coherent to doubt and reject any naturalistic or social definition of 'woman'. It is coherent to doubt any definition of goodness in terms of pleasure, desire-satisfaction, or any other natural and social features; e.g. it is coherent to doubt that for something to be good is for it to be conducive to pleasure. Similarly, it seems coherent—that is not self-contradictory—to reject an account of what it is to be a woman in terms of primary/secondary sex characteristics, chromosomes, socialization, gender identity, social subordination, or social constraints and enablements. It is coherent to say: 'Alex doesn't have female sex characteristics but, nonetheless, she's a woman given her gender identity' and it's also coherent to say, 'Alex does have a female gender identity, but nonetheless, she's not a woman because she doesn't have female sex characteristics'; we might think that someone who says this is mistaken, but someone who says this does not seem to be saying something incoherent or contradictory.[1] Indeed, many feminists argue that any attempt to reductively define what it is to be a woman will fail.[2]

These similarities between moral thought and talk and gender thought and talk give us reason to investigate the idea that, like moral thought and talk, gender thought and talk is inherently normative. In this paper I propose the following normative account of gender thought and talk:

> *The Fitting Treatment Account (FTA).* To judge that A is gender G (e.g. a woman) is just to judge that it is fitting to treat A as a G; and A is gender G is true iff it is fitting to treat A as a G.

FTA is a descriptive or hermeneutical account of our current gender judgments and gender concepts rather than an ameliorative account of the gender concepts we should accept (as, for instance, Haslanger's and Jenkins' accounts of gender are) or a metaphysical account of gender properties in social metaphysics (as, for instance, Ásta's account of gender categories is).[3] If we accept FTA, it is always possible to doubt a definition of 'woman' in terms of natural or social features F because for any set of natural or social features F it is always possible to judge that S has F but that it is not fitting to treat S as a woman. And FTA also explains why gender disagreements outstrip descriptive agreement: this is because disagreements about fit can outstrip descriptive disagreement. For instance, disagreements about the type of society it is fitting to desire and

[1] See McGrath (2021).

[2] See Mikkola (2019: §3) for an overview and discussion.

[3] See Haslanger (2012: ch. 7), Jenkins (2016), and Ásta (2018).

promote can outstrip agreement on the descriptive features of different societies; right-libertarians and socialists can agree about the descriptive features of societies governed by socialist and libertarian laws but still disagree about their desirability, and so about which society it is fitting to desire.

As I discuss in §4, at least one other *normative* account of gender which is a descriptive account of our gender judgments and concepts has been proposed.[4] One distinctive feature of FTA lies in the type of normativity that it understands gender judgments to involve. This normativity is not moral normativity, such as judgments about the gender one morally ought to treat someone as, nor is it all-things-considered normativity such as judgments about the gender one ought, all-things-considered, treat someone as. Instead, FTA understands gender judgments in terms of judgments about the gender it is fitting to treat someone as. What it is fitting for one to do and what one morally and all-things-considered ought to do come apart. For instance, in a much-discussed wrong kind of reason case a demon will murder everyone unless you admire it.[5] You all-things-considered and morally ought to admire the demon. But it is not fitting to admire the demon, since the demon is not admirable, and fitting admiration is admiration of the admirable. As I discuss in §4 and §5, the fact that FTA understands the normativity of gender thought and talk in terms of fit allows it to avoid wrong-kind-of-reason style problems that undermine Diaz-Leon's (2016) normative account of gender thought and talk and McGrath's (2021) expressivist account of gender thought and talk. McGrath's view makes it impossible for S to plan to treat A as a woman for moral or prudential reasons even though they do not judge that A is a woman; I argue that it should not have this consequence and FTA both does not have this consequence and explains why it can be possible for S to form such a plan. Diaz-Leon's (2016) view implausibly implies that if there are overwhelming moral benefits to misgendering A, such as to treating A as a woman rather than another gender, then it is true that A is a woman; FTA does not have this implication.

The other descriptive accounts of gender thought and talk that have been proposed are contextualist, such as Saul's (2012) proposal, and multiple concept approaches, such as Bettcher's (2013) and Laskowski's (2020) accounts. In §3-4 I argue that there are problems with these views: they implausibly imply that some genuine disagreements about gender, such as between trans-inclusive feminists and gender critical feminists, are not genuine disagreements; and they may seem to make it impossible for, for instance, trans women's claims about gender to be correct in the sense they believe them to be. FTA does not encounter these problems. And in §6 I argue that FTA does not face any other problems that render it implausible. If my arguments are sound, there are good reasons to accept FTA over the views with which it competes, FTA explains more than these views and does not face the problems that they face, and there do not seem to be other problems with FTA that give us strong reason to reject it. In this case FTA merits serious consideration as a descriptive or hermeneutical account of our gender thought and talk.

[4] Namely Diaz-Leon's (2016) proposal. A reader may associate the idea that gender is normative with Butler (1990). Butler's account of gender is not normally understood as a hermeneutical account of our gender thought and talk, our gender judgments and concepts and the meaning of gender terms; it may be better understood as an account of the metaphysics of gender. For how to understand Butler's view and critiques of Butler's view see Ásta (2018: ch.3) and McKittrick (2015: 2578-2579).

[5] See e.g. Rabinowicz and Rønnow-Rasmussen (2004).

2 | THE FITTING TREATMENT ACCOUNT

Fittingness is a normative property of attitudes and actions; it is a relation that holds between a possible action or attitude and a feature of the world that makes that action or attitude *merited*, *appropriate*, or *correct*. Fittingness is most often explained as a property of attitudes: fitting admiration is admiration that is merited by someone or something's admirable qualities; fitting desire is desire of the desirable or that which is worth desiring, fitting envy is envy of the enviable or that which is worth envying, fitting jealousy is jealousy of that which merits such a reaction.[6] Actions can also be fitting or unfitting. It is fitting to praise the praiseworthy, for instance, just as it is fitting to admire the admirable. It is unfitting for us to blame people who are not blameworthy and to assert that they are blameworthy. If someone deserves a compliment or a rebuke—and one has standing to give it to them—it is fitting to compliment/rebuke them accordingly. Fitting actions are actions that are appropriate given the circumstances or are merited by the context or the person in the context.[7] Suppose that you're curating an exhibition of paintings. It's fitting to put the most beautiful painting in pride of place. And it is fitting to pursue desirable things such as friendship and knowledge.[8]

Although fittingness was a much-discussed concept in the late 19th and early 20th centuries, much of the recent philosophical discussion of fit has arisen in the context of the 'wrong kind of reason problem' for accounts of value in terms of reasons for pro-attitudes. In the most famous case, a demon will punish everyone if you do not admire it. It seems that we have most reason to admire the demon, but the demon is not valuable, and accounts of value in terms of reasons hold that for X to be valuable is for there to be most reason to have pro-attitudes towards X such as to admire or desire it.[9] But although there is most reason to admire the demon, it is not fitting to admire the demon, after all the demon is not admirable and fitting admiration is admiration of the admirable. So, one response to this kind of case has been to move from accounts of value in terms of reasons for pro-attitudes to fitting attitude accounts of value, according to which for X to be valuable is for it to be fitting to have pro-attitudes towards X.[10] This discussion allows us to see that fitting attitudes are not attitudes that we morally ought to have or all-things-considered ought to have; for in the demon case we morally ought to admire it and all-things-considered ought to admire it, since this will do great good at no cost. The same distinction can be made in terms of actions: suppose that the demon will harm you and your family unless you publicly praise it; the demon is not praiseworthy so it is not fitting to praise it. But nonetheless you morally and all-things-considered ought to praise it. So, fit-based reasons for actions and attitudes are distinct from other forms of normative reasons such as moral and prudential reasons, and what we morally and all-things-considered ought to do comes apart from what it is fitting for us to do.

The similarities between moral disagreements and gender disagreements and the coherence of doubting any non-normative definition of 'good' and 'woman' give us good reason to investigate whether judgments of gender are normative (§1). Some reasons to think of fittingness as the particular type of normativity involved in gender judgments will come out in the discussion of normative contextualism and expressivism in §4-5—to preview, there are problems with thinking that

[6] See e.g. Howard (2018).

[7] See *ibid*. 9 n. 2, Broad (1930/1956: 221), and Ross (1939/1949: 81-82).

[8] See Howard and Rowland (2022).

[9] See e.g. Rabinowicz and Rønnow-Rasmussen (2004).

[10] See e.g. Howard (2019).

the relevant kind of normativity inherent in gender judgments is moral or all-things-considered normativity. But several considerations make understanding gender thought and talk in terms of fit relatively natural even in the absence of the comparison with normative contextualism and expressivism. First, the idea that gender might be understood in terms of fittingness has been briefly discussed or at least alluded to in the literature. Ásta (2018: 91) briefly discusses the view that 'X is a woman iff X merits the response in S that X be treated in manner M'; Ásta does not argue that we should reject this view, only that it provides an account of something different from that which her account of social gender properties is an account of.[11]

Second, many trans and non-binary people explain their genders in terms of fit. Discussing their experience of discovering that they are non-binary, Eli Erlick (2019: 231) says, 'I had already figured out "he"— the pronoun assigned to me at birth—didn't fit me'. Suzannah Weiss (2018) says that '[g]rowing up, I never felt people were wrong when they called me a woman, but it felt like a label imposed on me rather than one that fit. Then, in college, I learned about non-binary identity, and that did fit'. Similarly, Sinclair Sexsmith (2019: 109) says that '"woman" has never fit me' and CK Combs (2019: 94) says that after exploring their gender they 'found two new labels that seemed to fit [them] even better than "butch"—genderqueer and nonbinary'. Discussing how post-transition she came to identify as a woman rather than as genderqueer, Julia Serano (2016: 226) says that '[e]xperiencing the world (and my own body) as female makes the word "woman" feel like a far better fit for me now. (Susan Stryker (2008: 21) even explains gender identities in terms of our 'subjective sense of fit (or lack of fit) with a particular gender category'). We can take experiences like these of how it is fitting to refer to us, and what gender category fits or does not fit us, as experiences of the gender that we see it as fitting for ourselves and others to treat us as.

Should we think of these judgments as judgments about fit in the normative sense? Or might these judgments be explained in terms of a non-normative sense of fit, such as the sense of 'fit' that we make judgments about when we judge that our clothes don't fit? One reason to think that these judgments of fit are normative rather than non-normative is that (i) trans and non-binary people are willing to bear large costs—in monetary terms and in terms of ostracization by their families or others—to be treated in ways that they take to be fitting, and (ii) when they are not treated in these ways they judge to be fitting—and are instead treated in ways they take to be unfitting—a large majority of trans and non-binary people experience harmful and negative experiences, namely, gender dysphoria. These features are shared by other normative judgments of fit. For instance, we are often willing to bear very large costs to avoid doing something that we judge to be blameworthy. And we often have negative harmful experiences—guilt and shame—if we do things that we judge to be blameworthy. Similarly, suppose that you work at an Amazon warehouse and Jeff Bezos is scheduled to visit. Your boss says that they will dock your pay and may even fire you if you don't praise Bezos on his trip around the warehouse. But you think that Bezos is awful and anything but praiseworthy. You might bear these costs in order to refrain from praising him and if you did praise him, despite judging that he is anything but praiseworthy, you might feel bad and like you have betrayed yourself or your values.

[11] Ásta's account is a descriptive account in social metaphysics of how people are in fact gendered in particular contexts, and this merited treatment account would not be a plausible account of such properties, since someone can be socially treated as a G even though they do not merit being treated as a G. (For instance, a trans woman could be socially treated as a man even though she merits treatment as a woman). Ásta instead considers a merited treatment account of gender to be an account of whom we should treat as a particular gender; she does not raise any qualms with this view. But FTA is a descriptive or hermeneutical account of our gender judgments and concepts rather than an account of whom we should treat as a particular gender or an account in descriptive social metaphysics. Diaz-Leon (2016) and McGrath (2021) have also proposed accounts of our gender concepts in terms of how we should, or plan to, treat others, as I discuss in §4-5.

I propose that

> *The Fitting Treatment Account (FTA).* To judge that A is gender G (e.g. a woman) is just to judge that it is fitting to treat A as a G; and A is gender G is true iff it is fitting to treat A as a G.

FTA is intended to be a descriptive or hermeneutical view of our current gender judgments and gender concepts; it is not an ameliorative account of the gender concepts that we should accept, like Haslanger's ameliorative account of gender, or a descriptive account of social gender practices and the property of being a member of a social gender kind—that is, a metaphysical account of how people are in fact socially treated and gendered—like Ásta's conferralist account of gender; rather FTA is an account of our concepts of who is what gender. So, FTA does not compete with these views. Instead, it competes with other descriptive accounts of our gender concepts and gender judgments such as contextualism and expressivism. According to FTA, to judge that Alexa is a woman is just to judge that it is fitting to treat her as a woman; for instance, that it is fitting to use female pronouns to refer to her and/or to permit her to use, or to invite her to, women-only spaces.

FTA is analogous to fitting-attitude and buck-passing accounts of value. These views hold that for X to be valuable is for there to be reasons to value X or for it to be fitting to value X, and then give an account of what it is to value X in terms of a set of pro-attitudes (admiration, desire, approval, praise, etc.).[12] Similarly, FTA analyses judgments about S's gender in terms of judgments about how it is fitting to gender S or treat S as a particular gender. What is it to treat S as gender G? Treating S as gender G involves categorizing them as a G. There are many different ways of categorizing S as gender G including using pronouns associated with Gs, inviting them to G-only events, permitting them to use G-only spaces and otherwise grouping S with people who are Gs. Referring to someone as 'he', 'sir', or 'one of the blokes' is a way of treating someone as a man. Inviting someone to a girls' night out or a club for only women is a way of treating them as a girl/woman, as is telling them to line-up with the girls, rather than the boys, at school.[13] Forbidding someone from using the women's restrooms is a way of treating them as a man, or as not a woman.[14] So, according to FTA, to judge that S is a woman it is necessary for one to judge that it is fitting to categorize S as a woman.

It can be useful to think of this view about what it is to treat S as gender G in terms of norms. There are lots of different types of norms associated with particular genders. Some of these norms are what we can call categorization norms, such as the norm, 'use she/her pronouns for women but not for men'. To treat S as gender G involves treating S in line with the categorization norms associated with Gs such as the aforementioned pronoun norm. To judge that S is gender G is to judge that it is fitting to apply norms associated with Gs to S where these must include categorization norms associated with Gs. Some people's judgment that S is a G, which amounts to their judgment that it is fitting to treat them as a G, might involve the judgment that it is fitting to apply norms beyond categorization norms to S. For instance, some conservative people's (or people from the early 20th century's) judgment that S is a woman might involve the judgment that it is fitting for S to be a stay-at-home mother and so that it is fitting to encourage her to be a stay-at-home

[12] See e.g. Jacobson (2011), Rowland (2019), Suikkanen (2009), Skorupski (2010), and Way (2013).

[13] See Iantaffi and Berker (2018: 54) and Violet (2018: 24).

[14] See Kapusta (2016: 505) and Dembroff and Wodak (2018: 376-377).

mom, in addition to the judgment that it is fitting to categorize S as a woman in various ways. Liberals who hold that we should reject all norms about what women should do *qua* women will only judge that it is fitting to categorize S as a woman when they judge that it is fitting to treat S as a woman. The commonality here is the judgment that it is fitting to categorize S as a woman. To judge that S is a woman is just to judge that it is fitting to treat S in line with norms associated with women where these norms must include categorization norms but could, for a particular judgment maker, involve further norms associated with women. (We might think of norms associated with particular genders as falling into a core and periphery where categorization norms fall into the core and other norms such as norms of appearance, grooming, leisure activities, etc., fall into the periphery).[15]

The importance of gender categorization and gender categorization norms for our genders and gender practices is clear from trans communities' practices and trans people's experiences and writing about gender. Pronouns have a very important place in contemporary trans-inclusive discourses, practices, and communities. In *How to Understand Your Gender*, Alex Iantaffi and Meg-John Barker (2018: 54) introduce the concept of gender identity through an example of gender categorisation:

> Can you remember one of the first times that you realised what your gender was? It's OK if you can't. It can be as simple as being divided into girls and boys in school to line up, or for an activity, and having an inner sense of where you belonged. That inner sense may or may not have matched what others expected of you.[16]

And many trans and non-binary people explain their genders by reference to their experiences of how they should categorize themselves and which categorization norms they ought to conform with: explaining her gender, Julia Serano (2016: 78) recounts her experience as a child that she was doing something wrong by using the boys' restrooms; similarly, explaining her early experiences of her gender, Mia Violet (2018: 24) recounts her experience that she should line up with the girls rather than the boys at school.

The use of pronouns and the direction of someone to line up with people of a particular gender involves direct gender categorization, but there are other less direct ways of categorizing someone as a particular gender too. For instance, in *Detransition, Baby* one of trans author Torrey Peters' trans protagonists, Reese, remembers back to her childhood, in which she was assumed to be a boy. Her neighbour, Virginia, took her daughter and Reese to figure-skating lessons and Virginia began to (indirectly) categorise Reese as a girl in certain ways:

[15] Thanks to Carolina Flores and Elise Woodard for encouraging me to think of this distinction in terms of a core and periphery. To clarify, the idea that I have been articulating is *only* the idea that all that is common to all instances of someone's judging that it is fitting to treat S as a woman is the judgment that it is fitting to categorise S as a woman (in various ways) in virtue of certain features of her. It is consistent with this that we might take it as fitting to categorize S as a G (e.g. to use she/her pronouns to refer to her) in virtue of other features of them such as their sex characteristics, their social position, or their identity. Similarly, according to a popular account of wrongness all that is common to all judgments that A's ϕ-ing was wrong is that these judgments all involve the judgment that it is fitting to blame A for their ϕ-ing. But it is consistent with this view that we often judge that A's ϕ-ing is wrong (that is, fitting to blame) in virtue of other features of this action (e.g. that their ϕ-ing involved using someone as a mere means to an end, was not a utility maximizing action, or involved a breach of the set of moral rules that makes things go best).

[16] This is from a text box with the heading: 'Reflection point: Your gender identity'.

> …The way that Virginia included her with the other girls, complimented her on her grace, her form, the same as she did the others, so that eventually her daughter accepted Reese as one of them, and soon all of her friends did as well. The first time that Virginia just "forgot" and ordered five girl Happy Meals, instead of four girl Happy Meals and one boy Happy Meal.[17]

These indirect ways of categorizing someone as a particular gender are context-sensitive, can be context-specific, and can and do change over time.[18] We can, in general, treat S as gender G by applying norms other than direct and distinct categorization norms associated with Gs to them. For instance, in our current social context, by telling A to smile more, expressing shock at A's unshaven legs, or trying to convince A to wear a dress to their prom, A's mother treats A as a woman; by making fun of B for wearing makeup or telling them not to, B's father treats them as a man. But, like many feminists, we might reject all of these gendered norms of appearance and hold that it is unfitting to apply these norms to people—to hold that some people, in virtue of their gender, ought not have unshaven legs for instance. But if we do reject all these norms, we can still treat people as particular genders and judge that they are particular genders by categorizing them as particular genders, such as by referring to them with particular pronouns or permitting them to use particular single-gender spaces, and thereby judge that they are particular genders.

So, to treat someone as gender G is (at least) to categorize them as a G (directly or indirectly). And according to FTA, A is gender G is true iff it is fitting to treat A as a G, that is, iff it is fitting to (at least) categorize them as a G.[19]

Perhaps the currently most popular alternative to FTA, and so the most popular descriptive account of gender concepts and judgments, is contextualism. One of the key reasons to accept descriptive contextualism about our gender concepts is that it can explain why the way that we use gender terms seems to differ: some people seem to use gender terms to refer to sex; some people seem to use gender terms to refer to people's gender identities; others seem to use them to refer to the social roles that people occupy.[20] FTA also explains this: some people think that it is fitting, appropriate, or correct to treat someone as, and refer to someone as, a woman if and only if they have XX chromosomes or female primary sex characteristics; others think that it is fitting to treat those who have a female gender identity as women, and that it is unfitting and

[17] Peters (2021: 176)

[18] As can more direct ways of categorizing someone as a gender, as we will discuss in §6.

[19] The fact that there are both direct and indirect ways of treating and categorizing someone as a gender means that there are some complicated cases in which it is not entirely clear whether someone treats another as a particular gender. Suppose that Alice is Billy's mother. Alice treats Billy as a girl/woman in the sense that she tells Billy to wear a dress to their prom, tells them to smile more, and tells them to shave their legs. But Alice respects Billy's non-binary identity, or claims to, and their desires to be referred to with they/them pronouns and not to be referred to as one of the girls: she refers to Billy as 'they' and she doesn't talk about them as 'a girl'. Does Alice treat Billy as a girl/woman? It is tempting to say that she does. Although she does not explicitly categorize Billy as a girl/woman, it seems that Alice is only paying lip service to Billy's non-binary identity, and she is still treating them as a girl/woman. It seems that she is still categorizing them as a girl, albeit not explicitly, by applying these other norms associated with women to them. This case shows that although S is a G iff it is fitting to categorize them as a G, if one does enough things, which need not involve explicit categorization of S as a not-G, but do pragmatically imply this, one can fail to genuinely categorize S as a G—or can fail to categorize S as a G over time—even though one explicitly categorizes S as a G in particular utterances. (There may also be cases in which A treats B as gender G in certain respects or dimensions but not in others. For instance, suppose that B is a trans woman, A uses she/her pronouns to refer to B and permits B to use the women's bathrooms but, although she is a close friend of B's, A never thinks to invite her to girls nights' out.)

[20] See Diaz-Leon (2016: 245-246), Laskowski (2020: 40), and Saul (2012: esp. 196, 200).

inappropriate to treat these people as men, or not as women; others think that it is fitting to treat someone as a woman iff they occupy the social role of a woman. FTA also explains why, as I discussed in §1, it seems coherent to doubt any proposed definition of 'woman' in terms of natural or social features F. This is because, for any set of natural or social features F, it is always possible to judge that S has F, but that it is not fitting to treat S as a woman: it is possible to judge that S is socially positioned as a woman but that it is not fitting to treat S as a woman or that S has as female gender identity but it is not fitting to treat S as a woman.[21] FTA also explains why gender disagreements outstrip descriptive agreement. According to FTA, gender disagreements between gender critical feminists and trans-inclusive feminists about whether trans women are women outstrip descriptive agreement amongst the parties to this disagreement because even when they agree about all the descriptive non-normative facts, gender critical feminists and trans-inclusive feminists still disagree about whether it is fitting or correct to treat trans women as women.

§6 discusses several objections to FTA. But I want to deal with perhaps the most persistent here. This objection is that FTA is problematically circular because gender G occurs in the *analysandum* and the *analysans* of FTA. We can avoid this circularity by appealing to the norms which are socially associated with a particular gender.[22] According to social position accounts of the metaphysics of gender, all there is to gender is particular sets of norms, the endorsement of these norms, and the effects and experiences of these norms and endorsements of these norms in particular contexts.[23] Inspired by these accounts, we can say that to treat someone as gender G in context C is to treat them in line with a subset of norms, including categorization norms, associated with gender G or to treat these norms as applying to them. So, to treat A as a woman is to use she/her pronouns to refer to her, to invite, or permit, her to use women-only spaces, and to otherwise categorize her as a G (directly or indirectly) using norms associated with Gs in one's context. So, to judge that A is a woman is to judge that it is fitting to treat A in line with some of the norms, including categorization norms, associated with women or to treat (some of) the norms associated with women (including the categorization norms associated with women) as applying to A.[24] This doesn't yet get us out of circularity. But we know how to get out of the circularity now. We need an account of what distinguishes the (categorization) norms associated with women from the (categorization) norms associated with other genders and other social categories. And work on the metaphysics of gender—rather than on gender thought and talk—can help us here. For instance, we might take a Haslangerian approach to distinguishing gender and genders according to which: the norms associated with women are the norms associated with those who are marked as a target for subordination on the basis of actual or imagined bodily features presumed to be evidence of a female's role in biological reproduction.[25] If we combine this view with FTA, we hold that

> To judge that A is a woman is to judge that it is fitting to treat some of the norms, including categorization norms, associated with those who are marked as a target

[21] We might think of this as the open-question argument for FTA.

[22] For a slightly different response to a similar problem see McGrath (2021: 36-37).

[23] See, for instance, Ásta (2018), Barnes (2020) Dembroff (2018), Haslanger (2012), and Jenkins (2016) (2018).

[24] Similarly, on Jenkins' (2018) account of gender identity, to have gender identity G is to experience the norms associated with Gs as applying to one. This understanding of what it is to treat someone as a woman fits with Williams' (2016) discussion of what it is to treat an X as an F in general. Williams (*ibid*. 288) says that, '[t]here may be a certain suite of behaviours—certain *scripts*—such that *x* is treated as fragile [for instance] iff those scripts are applied to *x*'.

[25] See also Barnes (2020).

> for subordination on the basis of actual or imagined bodily features presumed to be evidence of a female's role in biological reproduction to apply to A.

This doesn't imply that in order to judge that A is a woman A must be marked as a target for subordination on the basis of such features or that A has such features.[26] We can judge that it is fitting to treat the norms associated with social group S to apply to A even if A is not socially treated as or understood to be a member of S; just as we can hold that it is fitting to treat A as a member of S even if A is not understood to be a member of S by most in one's society. For instance, we might think that it is fitting to treat A as cool even though they are not treated as cool by people in one's context or that it is fitting to treat A as queer, or part of the queer community, because they are bisexual even though bisexuals are not treated as queer in one's biphobic society.

It might be that this particular Haslangerian way of distinguishing the norms associated with women from norms associated with other genders is not quite correct; for instance it might be that we can instead distinguish the norms associated with women from the norms associated with men by holding that the norms associated with women are just those associated with those possessing all of the biological markers of female sex.[27] The important point is that we now have a recipe for providing a non-circular account of gender judgments and concepts. We do, or use, the work done by feminist philosophers and philosophers of gender to ascertain how to distinguish the norms associated with women from the norms associated with men, and then we plug that distinction into FTA.[28] The main thought behind FTA is that there are benefits to understanding gender judgments as normative, and understanding this normativity in terms of fittingness.[29] We can now see how to secure these benefits without circularity.

3 | ADVANTAGES OVER CONTEXTUALISM

In addition to being attractive in itself, existing descriptive accounts of gender thought and talk face problems, and these problems do not carry over to trouble FTA, as I'll argue in the next 3 sections. Perhaps the most currently discussed descriptive or hermeneutical view of our gender concepts and gender judgments is contextualism.[30] Contextualists hold that the meaning of 'A is a woman' varies depending on the context. For instance, according to the contextualist view that Saul (2012: 201) proposes, the meaning of 'woman' is determined by the standards operative in the context in which 'woman' is used. More specifically,

> X is a *woman* is true in a context C iff X is human and relevantly similar (according to the standards at work in C) to most of those possessing all of the biological markers of female sex.

One of the main motivations for contextualism is that some speakers appear to use gender terms like 'woman' as a sex term whilst other speakers seem to use 'woman' to refer to those

[26] So, this view does not encounter the inclusion problem that Haslanger's view faces; cf. Jenkins (2016).

[27] Cf. Saul (2012: 201)

[28] McGrath (2021: 37)

[29] *Ibid*.

[30] See e.g. Saul (2012), Diaz-Leon (2016), Laskowski (2020); see also Barnes (2020), and Dembroff (2021: 987).

who occupy a particular social gender role or identify as a woman. Yet these groups seem to be able to effectively communicate with one another about who is a woman.[31] How can our gender terms and our gender concepts be so flexible? As I explained in §2, FTA can explain this for it can hold that some speakers and communities think that it is fitting to treat those, and only those, with female sex characteristics as women; other speakers and communities think that it is fitting to treat those who have a female gender identity as women. Contextualist views like Saul's also explain the flexibility of gender terms and concepts. This is because in some conversational contexts (e.g. conservative communities) the operative standard for whether someone is a woman is whether they have XX chromosomes or whether they have female primary sex characteristics. But in other contexts (e.g. trans-inclusive communities) the operative standard for whether someone is a woman is whether they self-identify as a woman.

However, contextualism faces problems. Saul (2012: 209–210) and Diaz-Leon (2016: 247–248) argue that contextualism trivializes trans women's claims to womanhood because: although contextualism renders the claim that 'trans women are women' true, since this is true relative to the standards of trans communities and trans-inclusive contexts, it also renders the claim 'trans women are not women' true, since this is true relative to trans-exclusionary standards, such as the standards at play in a meeting of gender-critical feminists.

Consider a community where most people believe that to be a woman is to have XX chromosomes and that trans women should not be treated as women for practical and legal purposes: most members of this community believe that trans women should not be allowed to compete in women's sporting events, and should not be allowed to use women's restrooms or locker rooms. Contextualism implies that in the context of this community, it is false that trans women are women because trans women are not relevantly similar to most females according to the standards at work in that context.[32] According to Saul (2016: 210), this result

> [w]ould be deeply unsatisfying to the trans woman who wants to be recognized as a woman simply because *she is a woman* rather than because "woman" is such a flexible term. What the trans woman needs to do justice to her claim is surely not just the acknowledgment that her claim is true but also the acknowledgment that her opponent's claim is false. And the contextualist view does not offer that.

In contrast, FTA allows us to hold that trans women's claims are true and their opponent's claims are false: we can hold that trans women are women because we hold the first-order normative view that it is fitting to treat trans women as women; and so we can hold that people who deny this are mistaken that it is not fitting to treat trans women as women. But we can hold this view whilst simultaneously explaining why 'woman' can be used in many different ways. If we accept FTA, we can hold that 'woman' can be used in so many different ways for the same reason that 'just' can be used in so many different ways: different people have different conceptions of who it is fitting to treat as a woman just as different people have different conceptions of justice. We can hold that people who think that trans women are not women have a mistaken conception of who it is fitting to treat as a woman just as we might hold that right-libertarians have a mistaken conception of which societies are just.

[31] See Diaz-Leon (2016: 245-246), Laskowski (2020: 40), and Saul (2012: esp. 196, 200).

[32] *Ibid*. 247

Contextualism also struggles to explain disagreement about gender as well as FTA. Suppose that we are in a traditional society in which the vast majority hold a trans-exclusionary standard: suppose they hold that someone is a woman only if they have XX chromosomes, and we know this. Discussing Alice whom we know has XY chromosomes but identifies as a woman we say, 'Alice is a woman'. And the traditionalists say, 'Alice is not a woman'. We seem to disagree. And we seem to not be saying something obviously false. But contextualism struggles to accommodate this. If we and the traditionalists genuinely disagree, then we are mistaken and making a false claim, since according to the traditionalists' standard, which is the standard relevant in the context, Alice is not a woman. A proponent of this form of contextualism may say that the claim that we are making is false, but we just do not know what we mean here; we don't realise that we are making a claim, which if we understood its meaning, we would take to be obviously false. However, this seems implausible.[33] It seems that when we claim that Alice is a woman, we are not making a claim about the traditionalists' standards. Relatedly, unlike FTA, contextualism struggles to explain why it is always coherent to doubt any proposed definition of 'woman' in terms of particular natural or social features. For if a definition conforms to the standards at play in your context, and you know this, then according to contextualism, it should not be possible to coherently reject that definition. But it is perfectly coherent to reject traditionalists' definition of 'woman' in their context.

We might tinker with contextualism to try to make it explain disagreement better. We might hold that when S says 'A is a woman' the relevant standard for determining whether this claim is true is S's standard rather than the standard relevant in the context in which S utters this claim or in the context in which someone assesses this claim of S's. But this view generates the result that we and the traditionalists are not engaged in a genuine disagreement. For, on this view, when they say that 'Alice is a woman' their claim amounts to the claim that according to their standards she is not a woman and when we say that 'Alice is a woman' our claim amounts to the claim that, according to our standards, she is a woman. On this view, there is no proposition about which we disagree. In sum, contextualism seems to struggle to establish that we can always agree about all the relevant social and biological facts but still genuinely disagree about gender. But FTA explains this.

Contextualists about normative thought and talk have developed several responses to analogous problems concerning disagreement including responses that understand such disagreements as metalinguistic negotiations. Assessing whether possible replies to my argument along analogous lines might succeed would take an entire paper itself. However, very briefly, according to a metalinguistic negotiation response, we can understand the disagreement between gender-critical feminists and trans-inclusive feminists as just a disagreement about which gender concepts we should accept or as a disagreement just because both parties are advocating for rival ways of using 'woman'. It may be plausible to understand some such disagreements in these ways—that is, as metalinguistic negotiations. However, many trans women are not just advocating for a concept of woman on which they are women when they disagree with gender-critical feminists; many trans women also hold that *they are* women and that gender-critical feminists are mistaken about who are women.[34]

[33] McGrath (2021: 34-35)

[34] McGrath (2021: §2.1) argues against the metalinguistic negotiation response in more detail. I have also discussed problems with this and other normative contextualist responses to contextualists' problems with disagreement elsewhere; see Rowland (2020: ch. 3).

4 | ADVANTAGES OVER POLYSEMIC AND NORMATIVE CONTEXTUALIST VIEWS

In response to problems with contextualism, two types of views have been proposed that are similar to the contextualism we have been discussing. First, Talia Bettcher (2013) and Nick Laskowski (2020: 46–49) have proposed accounts on which there are multiple concepts of woman and man and gender terms are polysemous, that is, they can be used to express different concepts. On this view, those who hold that trans women are not women because they do not have XX chromosomes or were not born with female primary sex characteristics are using one (sex-based) concept of woman—and gender concepts more generally—and those who hold that trans women are women because they self-identify as women are using another (gender identity-based) concept of woman. As Laskowski (2020: 49) notes, it is unclear whether views like this count as contextualist or not, but they are at least similar to contextualism. On these views, we should not say that people who use trans-exclusionary concepts of gender are mistaken when they say that trans women are not women; rather we should say that they are using gender concepts that are morally objectionable because they are exclusionary, and we should advocate for inclusionary gender identity-based gender concepts.

I am not convinced that this view is an improvement on contextualism. Like contextualism, this view struggles to explain how we can disagree with those in trans-exclusionary societies about who is a woman. In fact, it seems to do worse here than contextualism. For it implies that when gender-critical feminists say that trans women are not women and trans-inclusive feminists say that trans women are women there is no hint of disagreement, since they are using different concepts of 'woman', and so they are talking past one another. Laskowski (*ibid*.: 47) argues that the talking past one another here is not the pernicious kind of talking past one another that occurs when two people argue about whether there is a bank in the town and one of them is talking about a financial bank whilst the other is talking about a riverbank. Laskowski says that we might think of the different concepts denoted by 'woman' as being more like the different related concepts denoted by 'cut' rather than like the different unrelated concepts denoted by 'bank': '"cut" has a meaning involving incision, another involving turning sharply, passing objectionably in a que, insulting someone, and so on'. However, suppose that Alex and Bec are in a disagreement about when Cath cut Dillon. Alex says Cath did cut Dillon, and Alex has in mind by 'cut' an insult, they mean that Cath insulted Dillon. Bec says that Cath did not cut Dillon, where Bec has in mind by 'cut' stabbing, they mean that Cath did not stab Dillon. Although there are similarities in the meaning of 'cut', Alex and Bec are still talking past one another in this case: they are not engaged in a genuine rather than merely verbal disagreement, since they do not disagree about what Cath did. So, Laskowski's and Bettcher's views imply that we trans-inclusive feminists cannot engage in a genuine disagreement with trans-exclusionary feminists, that when we seem to be disagreeing with them about who is a woman we are really just talking past one another. This does not seem right.

So, there are vices to the multiple concepts proposal. Furthermore, if we hold FTA, we can say everything that Bettcher and Laskowski want to say in different terms. They want to say that (i) we need to recognize that gender-critical feminists and trans-exclusionary people and societies use gender terms as sex terms and (ii) we need to campaign for trans-inclusive concepts of gender. If we accept FTA, we can say that (i) gender-critical feminists and trans-exclusionary people and societies have a sex-based *conception* of gender and particular genders such as 'woman', that is, they have a sex-based conception of who it is fitting to treat as a woman. But (ii) we should campaign for a trans-inclusive conception of gender and particular genders such as 'woman', and so of

whom it is fitting to treat as a woman. Similarly, socialists might say that in order to recognize the ills of contemporary society we need to recognize that most people hold capitalist or neo-liberal conceptions of justice and equality, and in order to achieve justice we need to campaign for the adoption of socialist conceptions of justice and equality.

Diaz-Leon (2016) has proposed a second normative response to problems with contextualism. She argues that we can keep contextualism but avoid its problematic implication that many people truly say that trans women are not women. We can do this by holding that the contextually relevant standard is always determined by the contextually relevant moral and political considerations. On this view,

> "X is a woman" is true iff X is human and relevantly similar to most females, where what counts as relevantly similar to most females depends on "objective" features of X's context, including instrumental, moral, and political considerations having to do with how X should be treated (regardless of who utters the sentence or what their beliefs are).[35]

Consider the context of a state legislating bathroom bills that outlaw trans women from entering women's restrooms. A state legislator says, 'trans women are not women'. Diaz-Leon's idea is that her normative contextualism implies that this claim is mistaken because practical, moral, and political considerations dictate that we should use a gender-identity-based standard for gender in this context. And according to this standard trans women are women. Note that Diaz-Leon (2016: 248–249) makes clear that her view is intended to be a descriptive or hermeneutical account of our gender thought and talk, and this means that it must be giving an account of our current gender concepts and the judgments about gender that we currently make.[36]

However, as a descriptive rather than ameliorative view of our gender judgments and concepts, it seems that Diaz-Leon's view puts the wrong kind of normativity into our gender judgments and concepts. Suppose that Amy judges that (i) trans women are not women because she holds a chromosomes or socialization-based conception of who is a woman. But Amy thinks that (ii) she morally and practically ought to treat trans women as women because she does not see any downside to using she/her to refer to trans women or to trans women using women-only spaces. And she does see a downside to not doing this: she knows her failing to treat trans women as women will harm trans women in a variety of ways; e.g. by refraining from treating trans women as women she will cause trans women harmful social gender dysphoria.[37] Diaz-Leon's view makes Amy and her judgments impossible and incoherent But Amy's position and judgments seem entirely intelligible.[38]

It might be responded that Amy probably believes both that

A. The standard for who is a woman that is morally and politically most useful in her context is not the self-identification (SI) standard on which trans women are women; and that
B. For moral and political reasons, she should nonetheless treat trans women as women.

[35] Diaz-Leon (2016: 251)

[36] See also Laskowski (2019: 45-46).

[37] See e.g. Kapusta (2016) and Badgley et al (2021).

[38] I do not mean to condone this position but only its intelligibility.

If Amy believes both (A) and (B), then Amy's judgments are entirely consistent with Diaz-Leon's view, since on Diaz-Leon's view to judge (A) is to judge that not all trans women are women.

However, I am imagining that Amy does not believe (A); I am imagining that she rather believes that the SI standard is the morally and politically most useful standard in her context because adopting this standard for the categorization of who is a woman allows us to effectively mitigate trans women's gender dysphoria and given that Amy judges that there are no downsides to refraining from treating trans women as women she does not judge that there are any moral and political considerations that make it morally and politically better to use another alternative standard; and so no considerations make any such alternative standard salient. Amy does not judge that categorizing trans women as women harms the interests of cis women or that we need to use a standard other than the self-identification standard in order to fully understand or articulate the oppression that women face. Amy just thinks that womanhood is based on socialization or sex/chromosomes; perhaps she has read Theodore Bach's (2012) work and on this basis is convinced that to be a woman is to have been socialized as a woman, or perhaps she just cannot shake a sex-based or chromosomes-based conception of gender. This version of Amy who holds that the SI standard is the most morally and politically useful—that is, who holds the negation of (A)—but holds that trans women are not really women still seems entirely intelligible, and Diaz-Leon's account implies that this version of Amy is unintelligible.[39] Perhaps most Amy's will hold (A), but some will and do hold the negation of (A).[40] In this case, Diaz-Leon's account implies that gender judgments that are intelligible are not intelligible and a descriptive, rather than ameliorative, account of gender concepts and judgments should not imply this.

In contrast, FTA is compatible with the view that, and may even predict that, people like Amy will be possible. Often what we judge that it is fitting for us to do is different from what we judge that it would be best for us to do. If an evil demon will punish us unless we admire it, it is not fitting for us to admire the demon, for they are not admirable, but it would be best for us to admire the demon. Suppose that you are the curator at the Met. A terrorist threatens to harm your family unless you hang a painting they made in pride of place at the latest exhibition. It would be best to hang the terrorist's painting in pride of place, you probably should do this, and you may do this, but it is not fitting to hang it there, and you would not judge it to be so fitting to do so, since the painting does not have features that make it worthy of being so esteemed. Or suppose that a dictator will kill many people unless you frame innocent Isla for a crime, blame her, and assert that she is blameworthy. It may be best to blame Isla and to assert that she is blameworthy. But it is not fitting to blame her and nor is it fitting to say that she is blameworthy, since she is not. And you may judge that you morally ought to blame Isla or assert that she is blameworthy but that she is nonetheless not blameworthy. Similarly, FTA explains why Amy can simultaneously judge that trans women are not women and that morally she ought to treat trans women as women. This is because according to FTA, when Amy judges that trans women are not women she is judging that it is not fitting to treat trans women as women, because they do not have features such that it is fitting to treat them as women on her conception of who it is fitting to treat as a woman. And

[39] Similarly, if an all-powerful evil dictator will punish everyone in our society whenever someone does not praise them, we can hold that the morally and politically most useful and salient standard for who is praiseworthy in that context include features that the dictator has. But we can simultaneously hold that the dictator is not praiseworthy, praise does not fit them because the salient standard and most morally and politically useful standard for praiseworthiness is incorrect; for it is not fitting to praise the dictator, praise doesn't fit them and they are not worthy of it.

[40] Amy's position seems similar to the position that some people who do not think that trans women are women, such as British MP Jess Phillips, take, see e.g. Turner (2021).

one can think this whilst thinking that morally and politically one ought to treat trans women as women and that the relevant morally salient standard for womanhood in your context implies that all trans women are women.

Similarly, suppose that Charlie is non-binary and detests being treated as a woman. But Charlie was assigned female at birth, they do have some female sex characteristics and do sometimes express themselves in a feminine way. Now imagine that a dictator will kill everyone in our large society unless we treat Charlie as a woman for the next week, refer to them as a woman, invite them to women-only events and spaces and apply stereotypes associated with women to them. It will follow from Diaz-Leon's view that it is true that Charlie is a woman in this context. But this seems wrong. The dictator's threat seems irrelevant to Charlie's gender.[41] In contrast, according to the fitting treatment account, the dictator's threat is irrelevant to Charlie's gender, since the dictator does not make it the case that it is fitting to treat Charlie as a woman but only that there is most moral and political reason to do so.

I take the upshot of these problems for Diaz-Leon's view to be that although Diaz-Leon is on to something in thinking that our gender judgments are normative, she puts the wrong type of normativity into our gender concepts. The relevant type of normativity is not all-things considered practical normativity, or moral and political normativity, but is rather fittingness. That is, the reasons we take to be relevant to determining whether A is gender G are just the fittingness-based reasons bearing on whether we should treat A as a G—the reasons that determine whether it is fitting to treat A as a G—rather than all of the reasons bearing on whether we should treat A as a G such as all the moral and political reasons bearing on whether we should treat A as a G.[42]

5 | ADVANTAGES OVER EXPRESSIVISM

Sarah McGrath (2021) has recently proposed an expressivist account of gender according to which to judge that someone is gender G is to plan to treat them as a G. And 'X is a woman' means 'I plan to treat X as a woman' because to assert 'X is a woman' is to express one's plan to treat X as a woman.[43] Expressivism faces several problems which FTA does not face.

First, it seems possible to judge that someone is gender G without planning to treat them as a G. Consider again Amy who judges that trans women are not women but does plan to treat trans women as women because she does not see any downside to using she/her to refer to trans women or to trans women using women-only spaces, and she does see a downside to not doing this: she knows her failing to treat trans women as women will harm trans women in a variety of ways, such as causing the trans women she does not treat as women social gender dysphoria. When Amy judges that trans women are not women she does not plan to refrain from treating trans women as women. McGrath's expressivism seems to make Amy impossible.[44]

[41] If this case is not sufficiently counter-intuitive suppose that Beth is a cis woman and the dictator will kill everyone unless we treat her as a man for the next week. Diaz-Leon's view also seems to imply that it is false that Beth is a woman.

[42] For other different problems with Diaz-Leon's view see Laskowski (2019: 43-46) and McGrath (2021: 34-35).

[43] McGrath is not explicit about what her view implies for the meaning of sentences involving gender terms, but one of the core ideas behind expressivism as an approach to how to understand our thought and talk about sentences in domain D is that the meaning of a sentence S is determined by what it is to think S; see e.g. Schroeder (2010: 74). So, we should understand McGrath's expressivism as a view about the meaning of gender terms and sentences involving gender terms as well as about gender judgments.

[44] Furthermore, McGrath's expressivism implies that it should seem incoherent for Amy to express these judgments: just as it seems incoherent to sincerely say, 'it's raining but I don't believe that it's raining', if McGrath's expressivism were

One way to respond to this problem is to say that people like Amy plan to treat trans women as women in some ways or contexts but not in others. On this view, Amy must plan to not treat trans women as women in some ways or contexts.[45] However, we can imagine versions of this case where this is not so. We can easily imagine Amy intelligibly saying something like: 'okay, I'm only going to say this once because it doesn't really matter or affect what we do socially, and I'm not going to raise it when we talk about gender after this, but actually trans women are men…' before listing her reasons why. This seems like something someone could coherently and sincerely say. McGrath's expressivism makes this impossible. If McGrath's view were correct, Amy would be saying something contradictory. But what Amy says does not seem to involve a contradiction. And McGrath's expressivism is a descriptive or hermeneutical rather than revolutionary or ameliorative view of our gender thought and talk. So, whether there seems to be a contradiction here according to competent speakers matters; this view predicts that there should be such a seeming contradiction.

The problem here is similar to the problem with Diaz-Leon's account, not just in the example but in terms of what the source of this problem is: what we plan to do is often what we think we all-things-considered ought to do, or what we have settled on doing given all kinds of different considerations including moral and prudential considerations. But sometimes we can plan to treat someone as a particular gender due to there being very strong moral or prudential reasons for so treating them even when we do not judge that they are that gender; just as we can plan to treat someone as praiseworthy due to the moral or prudential reasons to do so when we do not believe that they are praiseworthy. In both cases, we may plan to do something—or judge that we ought to do something—that we do not judge that it is fitting or correct for us to do.

Might McGrath respond to this problem by arguing that Amy plans to treat trans women as women in the world in which there are no bad consequences to treating trans women as women? I don't think so. Suppose that Amy thinks that part of what it is for A to be a trans women is for A to have social gender dysphoria such that being treated as a man or not as a woman by others causes A significant discomfort, pain, or suffering; a not uncommon view.[46] *And* suppose that Amy thinks that in every world in which there are trans women, treating trans women as not-women, even in private, would have worse consequences than treating trans women as women. It still seems possible for Amy to coherently judge that trans women are not women but plan to treat trans women as women due to the bad consequences of not doing so—and to say as much. Yet in this version of the case Amy does not believe that there exists a possible world in which there are trans women and there are no bad consequences to treating trans women as men or not-women that would stop her from planning to treat trans women as women. So, in this case Amy does not plan to treat trans women as men or not-women in any possible world yet still coherently judges that trans women are not women.

correct, it should seem incoherent to sincerely say, 'trans women are not women but I plan to treat them as women'. For just as the sincere assertion, 'it's raining' expresses one's belief that it is raining, similarly, according to expressivism, the sincere assertion 'trans women are women' expresses one's plan to refrain from treating trans women as women. Yet it does not seem at all incoherent or Moore-paradoxical for Amy to say, 'trans women are not women but I plan to treat them as women'. Similarly, Diaz-Leon's account will imply that Amy's claim that 'trans women are not women but morally and politically I ought to treat trans women as women' (§3), is Moore-Paradoxical. But this claim does not seem to be Moore-Paradoxical. Cf. Woods (2014: 4)

[45] McGrath (2021: 36-37) (p.c.)

[46] Although we should reject this view; see e.g. Ashley and Ells (2018: 24).

Another response to this argument might seem plausible. Call the set of possible worlds in which there are no trans women PN and the set of possible worlds in which there are trans women PT. It could be claimed that Amy plans to treat trans women in (some) PT worlds as not-women if she is ever in a world in the PN set. So, according to this response, she has trans-world plans: if she finds herself in a world where there are no trans women around she plans to not treat merely possible trans women (that is, trans women in other possible worlds) as women. However, this response cannot succeed for Amy knows that she will never be in a PN world. So, she cannot make such a plan, since I cannot plan to ϕ in C if I know that I will never be in C. Furthermore, it seems implausible to think that the meaning of 'trans women are not women' is determined by a hypothetical plan about extremely remote worlds containing no trans women.

In contrast, as we discussed in §4, FTA is compatible with the view that, and may even predict that, people like Amy will be possible. This is because often what we judge that it is fitting for us to do is different from what we judge that it would be best for us to do; for instance, we might judge that it is not fitting for us to praise Bezos but we might plan to praise him and judge that it would be best for us to do so, because we will lose our job at the Amazon warehouse if we do not praise him (§2).[47]

A second problem for McGrath's expressivism about gender is a problem that expressivism in other domains also faces concerning embedded gender terms. Expressivist views about thought and talk in particular domains D (morals, modals, law) hold that D-language functions differently from descriptive language. According to expressivists, descriptive language (e.g. 'it is raining outside') expresses beliefs about the world but D-language expresses desire-like states such as plans. But for many domains D for which expressivism is proposed, D language functions like descriptive language. The relationship between descriptive claims like, 'it's raining outside' and related questions, conditionals, and negations—'is it raining outside?'; 'if it's raining outside, Alex is getting wet'; 'it's not raining outside'—is the same as the relationship between D claims and the analogous related questions, conditionals, and negations. For instance, 'stealing is wrong' has an analogous relationship to 'is stealing wrong?', 'if stealing is wrong, then stealing from Walmart is wrong', and 'stealing is not wrong'. The same is true of gender claims: 'Alex is a woman' bears an analogous relationship to 'is Alex a woman?', 'Alex is not a woman', and 'if Alex is a woman, then

[47] Doesn't FTA also imply that claims which do not seem Moore-paradoxical in fact are; see *supra* note 44. For instance, it implies that it is incoherent to say that

(*) 'Alice is a woman but it's not fitting to treat her as a woman'.

But it might be argued that (*) doesn't seem obviously incoherent. I do not think that the problem here is very strong for FTA in the way that there is a strong analogous problem for expressivism. First, in Amy's case, her saying, 'trans women are not women but I plan to treat them as women' is the only way for her to communicate her views about gender and her plans going forward. Expressivism implies that she cannot coherently do this. So, in a certain context expressivism renders a sentence incoherent which is the only way of expressing something very clear which cannot be otherwise clearly expressed. In contrast, there is no obvious context in which (*) might to be used to communicate something. Relatedly, if we rejected FTA just because (*) is not obviously incoherent, we would seem to be committed to the view that there are no interesting conceptual analyses, since all unobvious conceptual analyses will be analyses that it is not incoherent to deny; cf. Miller (2003: 16-17). But we are not committed to this view if we reject expressivism on the basis that I have been arguing against it, since expressivism does not just produce incoherence where it is not obvious that there is incoherence; expressivism makes sentences incoherent that have a clear meaning and that we might use to express something important. Finally, as I explain in §6, it is fitting to ϕ only if it is in some sense correct to ϕ. And it is not obviously coherent to claim, 'Alice is a woman but it's not correct to treat her as a woman'. Indeed this does seem incoherent, since by saying that Alice is a woman I appear to be saying that according to some standard, which I take to be salient, it is correct to hold that she is a woman; in this case it is correct that she is a woman and it is correct to treat her as a woman in the sense of taking her to be a woman according to that standard.

Beth is a woman'. Just as the predicate 'raining' seems to mean the same thing when the predicate is embedded in a conditional, negation, or question, similarly the predicate 'a woman' seems to maintain the same meaning when embedded in a question, negation, or conditional. Expressivists about morality—as well about other types of thought and talk such as about pejoratives—have struggled to secure the result that moral predicates preserve the same meaning when embedded in a negation, conditional, or question: expressivists have struggled to solve the Frege-Geach problem.[48] So, we should expect expressivism about gender to face a similar struggle. And McGrath does not discuss how expressivism about gender deals with gender terms in embedded contexts such as conditionals and negations.

It shouldn't be controversial that proponents of expressivism about gender at least need to show that expressivism can explain how gender terms function and maintain their meaning in embedded contexts, especially given that expressivists in other domains have faced such problems here. But to quickly get a slightly more specific grasp on this problem consider the question, 'is Reese a woman?' 'A woman' seems to have the same meaning in this question as it does in the claim 'Reese is a woman'. We want an account of the meaning of gender terms that can explain this and give a plausible account of what this question means. But it is hard to even understand what 'is Reese a woman?' means if we adopt expressivism. If when I say, 'Reese is a woman', this expresses my plan to treat Reese as a woman, what does it mean to ask, 'is Reese a woman?'? This question cannot plausibly express an internal question that I have about whether I plan to treat Reese as a woman. Perhaps McGrath's expressivism could be combined with an expressivist account of the meaning of normative terms. In this case questions about gender, such as 'is Reese a woman?', could be analysed in terms of questions about whether we *should* treat someone as a particular gender that get an expressivist analysis. But this would involve hitching McGrath's expressivism about gender to expressivism about normativity, which is extremely controversial; it would be better to hold a view about gender thought and talk that has fewer metanormative commitments. Furthermore, similar problems arise for expressivism about normative thought and talk of course. For instance, if we hold that to judge that we should ϕ is to plan to ϕ, then we need an account of what it is to ask the question, 'should I ϕ?' And, again, this cannot be plausibly held to express an internal question about whether we plan to ϕ.

In contrast, FTA does not face the Frege-Geach problem, since it does not hold that claims about gender express desire-like states; it is consistent with FTA that gender claims express beliefs about fitting treatment. And the Frege-Geach problem only arises for the meaning of sets of claims C according to which these claims express desire-like mental states or express mental states rather than their content. FTA holds that the meaning of gender claims consists in the content of the judgments of fit that they express. On this account, 'Reese is a woman' means 'it is fitting to treat Reese as a woman'.[49] So, on this account, 'is Reese a woman?' means 'is it fitting to treat Reese as a woman?' Perhaps the Frege-Geach problem for views like McGrath's expressivism can be overcome. But no solution to the Frege-Geach problem has achieved widespread acceptance and it is controversial whether any proposed solution succeeds. So, other things equal, it seems that we

[48] For discussion of this problem in general terms see Geach (1965) and Schroeder (2008: 5-6). For discussion of how expressivists about morality and slurs have struggled to deal with this problem see e.g. Schroeder (2010: ch.6-7) and Hom (2010).

[49] Or more specifically, if we take the Haslangerian view discussed in §2 for instance: it is fitting to treat A in line with the (categorization) norms associated with those who are marked as a target for subordination on the basis of actual or imagined bodily features presumed to be evidence of a female's role in biological reproduction (such as pronoun norms).

should prefer an account like FTA, which does not face the Frege-Geach problem, to an account like expressivism which does.

FTA has two further advantages over expressivism. To get the first in mind it is best to consider a problem that expressivism about slurs has been held to face. According to expressivist accounts of slurs, slurs express contempt via their meaning. On this view, when A says that B is a Kraut, they express contempt for B. But if this is all the negative attitudinal content that slurs like this one express, then it is unclear that this fully captures what's going on with slurs. Part of what is so destructive about slurs, it has been argued, is that they enjoin others to take similar attitudes of contempt to the slurred: if I say that B is a Kraut, I enjoin others to have contempt for B too (if this is hard to get in mind with 'Kraut', think of a more obviously offensive term). The diagnosis that many have been attracted to here is that when A says that B is a Kraut they do not just express contempt towards B, they represent B as *contemptible*, they say that B merits, or is a *fitting object of contempt*.[50] One idea here is that we need the idea that the slur represents the slurred as contemptible rather than just expressing contempt because if the slur just expressed contempt—if the slur just vented the speaker's feelings and did no more—then the contempt expressed could be written off as just a problem for the speaker. The problem is that slurs do more than vent the speaker's feelings; slurs seem to express feelings whilst expressing the view that these feelings are apt or appropriate. This is how slurs can enjoin others to have similar attitudes, since apt or appropriate feelings are feelings that it is appropriate for others to have too.[51] So we might think that a view on which when A says that B is a Kraut they say that (or pragmatically imply that) B is *contemptible* thereby has an advantage over an expressivist view that holds that when A says that B is a Kraut A merely expresses their contempt towards B.

FTA may have an analogous advantage over expressivism about gender. When someone misgenders another or when A denies B's gender identity by saying, 'you're not a woman, you're a man', we might (similarly) think that it does not fully capture what is going on here to say that A just expresses a plan to treat B as a man rather than as a woman. It seems that A does more than this; they say that it's *appropriate, fitting, or correct* to treat B as a man rather than a woman. When someone says that you are a man rather than a woman, they are not just expressing their idiosyncratic plans to treat you in a particular way but saying something about how they think it is appropriate, correct, or right to think of you and treat you. Furthermore, misgendering someone is often held to involve erasing their gender identity.[52] It is much easier to explain this if part of what it is to say that someone is a man is to say that it is appropriate to treat them as a man rather than just that one plans to treat them as a man. If one says that it is appropriate to treat A as a man knowing that A identifies as a woman, one takes their gender identity to be irrelevant to their gender, and the gender it is fitting to treat them as; this is less clear if one merely doesn't plan to treat them in line with their gender identity.

Finally, as we discussed in §2, many trans and non-binary people understand and explain their genders in terms of fittingness. FTA explains why this is. It is unclear how expressivism could explain why many trans and non-binary people understand and explain their genders in terms of fit. Perhaps the only way we could explain this if we accepted expressivism would be if we adopted an expressivist account of fittingness. But expressivist metanormative views, such as expressivist accounts of fittingness, reasons, and value, are extremely controversial. At the very least, FTA

[50] See e.g. Jeshion (2018: esp. 86-87), Bach (2018: 69), and Camp (2013: 333).

[51] *Supra* note 50.

[52] See Bettcher (2014: 392), Kapusta (2016: 509), and Dembroff and Wodak (2018: 383).

provides a straightforward explanation of why many trans and non-binary people understand and explain their genders in terms of fit; expressivism cannot provide this straightforward explanation and it seems difficult for expressivism to explain why trans and non-binary people understand their genders in this way.

6 | OBJECTIONS TO THE FITTING TREATMENT ACCOUNT

I've been arguing that there are reasons to accept FTA over other descriptive accounts of our gender concepts and gender judgments that have been proposed in the literature. In closing I want to consider some natural worries and objections to

> *The Fitting Treatment Account (FTA)*. To judge that A is gender G is just to judge that it is fitting to treat A as a G; and A is gender G is true iff it is fitting to treat A as a G.

One set of worries concerns subordination and oppressive norms. First, it might seem that FTA implies that when someone who was assigned male at birth judges that they are a woman they are judging that it is fitting for them to be subordinated. However, one can judge that it is fitting to treat A in line with (categorization) norms associated with subordinated group S without thereby thinking that it is fitting to subordinate A and whilst simultaneously thinking that we should not subordinate members of S. One can judge that it is fitting to refer to A with the pronouns associated with members of S, to permit A to use the spaces that only members of S are permitted to use, and to group A in with the women/girls rather than the boys/men when there is such a grouping without judging that it is fitting to subordinate A in any way. Relatedly, it might seem that FTA cannot be right because women in oppressive states, where women are forbidden from driving for instance, do not all think that it is fitting for them to be treated as the majority treat women. However, although such women do not think that it is fitting for women to be treated in the way *the majority treat women*, they still think that it is fitting for them to be *treated as women*: to be referred to with she/her pronouns and to be given access to female toilets and changing rooms and other women-only spaces; they do not think it is fitting for them to (only) be permitted to enter male-only spaces or to be referred to with he/him pronouns.

There are related issues about *when* it is fitting to treat someone as a particular gender. Suppose that we should abolish gender, abolish all norms associated with genders.[53] Does that mean that it is not fitting to treat anyone as having a gender now and so that, if FTA holds, no one has a gender now? FTA can stay relatively neutral here. But suppose that it is fitting *and* morally right to abolish gender. In this case it may not be fitting now to treat people as particular genders, and it would not be fitting to treat anyone as any gender if gender were abolished. But it is consistent with this that it is *conditionally* fitting to treat people as particular genders: if we assume a gender categorization scheme like that which we currently have, or *if we're going to treat people as having particular genders*, then it is fitting to treat everyone who is a man (or who has features—e.g. a male gender identity—that make it fitting to treat them as a man) as a man.[54] Furthermore, even

[53] See Cull (2019) for discussion.

[54] Similarly, suppose that you are a moral error theorist who believes that we should abolish morality. In that case, you can hold that, assuming our current moral categories, there are some actions that are morally wrong—such as pre-meditated murder—and there are some actions that are morally permissible and innocuous—such as brushing your hair. It is fitting to treat brushing your hair as permissible and murder as wrong given our current moral categories, but we should abolish

if it is both (i) morally right to abolish gender and (ii) fitting to bring about the abolition of gender this does not establish that (iii) it is not fitting to categorize or treat anyone as a particular gender now, when gender has not yet been abolished.

Another possibility is that, due to their gender identities, it is (unconditionally) fitting to treat people with male gender identities as men and those with female gender identities as women and because of this it is not fitting to abolish gender but that morally we ought to abolish gender and so morally we should not treat anyone as a particular gender.[55] In this case, we have a different view regarding (ii): it is not fitting to abolish gender completely but morally we ought to abolish gender. Of course, it could be fitting and morally right for us to abolish all gendered social roles, scripts, and norms of appearance associated with genders—to no longer have norms such as 'women should shave their body hair' and 'women should be slim but not muscular'—without it being fitting and morally right for us to abolish our practice of categorizing people as different genders. In this case it would be fitting to categorize some people as women and others as other genders even though it is not fitting to apply any pernicious gendered norms concerning social roles and appearance to anyone.

What about if what it is to treat or categorize someone as a particular gender changes? Suppose that at time T1 we use she/her to categorize people as women but at T2 we stop doing this and instead use zhe/zer pronouns for women (she/her pronouns come to be associated with another gender). Furthermore, at T2 we stop associating gracefulness with girls and women, and so trans girls and women—such as Torrey Peters' Reese (§2)—would not feel affirmed and treated (indirectly categorized) as girls/women if someone went out of their way to compliment them on their gracefulness; instead intelligence comes to be distinctively associated with women (and not with men). It seems to me that we should think of these changes of norms as having the followimg effects on fittingness. Suppose that Alyssa is a trans woman at T1. At T1, referring to her with she/her pronouns and going out of one's way to compliment her on her gracefulness is a way of treating her as a woman. It is fitting to treat Alyssa as a woman at T1 and at T2 (suppose). So, it is fitting to do these things at T1 but it is not fitting to do these things at T2, since doing these things at T2 would no longer involve treating Alyssa as a woman: at T2 it is fitting to refer to Alyssa with zhe/zer pronouns and to go out of one's way to compliment her on her intelligence instead.[56]

A different worry about FTA is that we learn gender concepts before we learn about fittingness: for instance, a 9-year-old can judge that some people are women but need not make any judgments about fittingness.

However, we do have the concepts picked out by philosophical talk of fittingness from an early age. We learn that certain forms of behaviour are inappropriate in particular places, that it is correct to put certain objects in particular places or to write in one way rather than another, and we learn that it is correct for some people to line up with the girls and correct for others to line up

these categories and so if we do not presuppose these categories, it is not fitting to treat these actions in these different ways; cf. Evers and Streumer (2016).

[55] Although if we are going to treat some people as a particular gender G, then we ought to treat everyone with gender identity G as a G.

[56] I'm assuming that Alyssa's view of the pronouns that fit her tracks those that are used to categorise someone as a woman in her society. Her view of the pronouns that fit her may not track this in two different ways. First, Alyssa may hold that using she/her to refer to her still involves categorizing someone as a woman at T2 even though socially this is not taken to be the case at T2; as some people hold that referring to oneself as 'queer' is still derogatory even though it is not socially held to be such anymore. Second, she may prefer to be referred to as she/her, just because she likes these pronouns, or she might have a (political) view that to be referred to with a particular pronoun is not to be categorized as a particular gender; he/him lesbians have a view about he/him pronouns along these lines.

with the boys when there is such a split.[57] So, it seems that we do make judgments of fit from an early age. Some of the plausibility of this response relies on the view that to judge that ϕ-ing is correct is to, or is one way of, judging that ϕ-ing is fitting. Fittingness and correctness are often used interchangeably.[58] For instance, according to Brentano's (1969: 18) fitting-attitude account of value, that which is valuable is that which it is fitting *or correct* to love. But even if fittingness and correctness are not exactly interchangeable, it is plausibly fitting to ϕ iff it is correct to ϕ. For instance, it is fitting to admire someone iff it is correct rather than incorrect to admire them.[59] There is another way of seeing that we make judgments of fit from an early age too: to see something as desirable is to see it as a fitting object of desire and we see things as desirable from an early age, since we desire things from an early age, and on many views this involves seeing things as desirable.[60]

Furthermore, we might question whether X can figure in a good conceptual analysis of Y or account of judgments of Y only if we become acquainted with or learn to make judgments of X prior to or at the same time as becoming acquainted with or making judgments of Y. Michael Smith (1997: 103) argues that X can provide a good conceptual analysis of Y so long as X best explains our inferential judgments about Y and the other ways in which we use Y. FTA may be the best explanation of our judgments of gender even if we make judgments of gender prior to making judgments of fit. For when we are young our judgments of gender may just parrot the gender divisions/judgments of those around us—as our moral judgments may parrot the moral judgments of those around us when we are young given that we lack the moral/conventional distinction at a young age. But the gender judgments of those older than us that we are parroting may still be judgments of fitting treatment. Furthermore, the arguments that I've been making in this paper seem to show that FTA explains our judgments of gender better than contextualist, polysemic, and expressivist views of gender thought and talk.

A final issue for FTA is that it might seem to be overly complex or to just not be particularly intuitive. When we say that someone is a woman it might seem that we are just not saying anything so complex as saying anything about how it is fitting to treat them, and it might just seem that we are not saying anything about how it is fitting to treat them at all. However, fitting treatment in FTA could be seen as a gloss on the simple and the intuitive idea, which we have seen that many trans and non-binary people pre-theoretically have (§2), that whether someone is gender G is a matter of whether G fits them, or whether ways of categorising them as a G seem to fit them, who they are, or their qualities. The idea that someone fits in with the other girls or that the label boy fits someone is not a very complex idea at all, nor is the idea that it is correct to say that someone is a woman or correct to use she/her pronouns to refer to her.[61] But even if it were the case that FTA is overly intellectual and not particularly intuitive, in §3-5 of this paper I've argued that there are strong reasons to accept it over the alternative views with which it competes; these strong reasons would outweigh any lack of intuitiveness or complexity which FTA may seem to have.

[57] See Iantaffi and Barker (2018: 54) and Violet (2018: 24).

[58] See e.g. Jacobson (2011; §1) and Howard (2018: esp. 9 n. 4).

[59] For an explanation of the relationship of fittingness to correctness see Howard and Rowland (2022).

[60] See, e.g. Gregory (2021: esp. ch. 6) and Shaw (forthcoming). Furthermore, to judge that X is desirable is to judge that it is good (in a certain way); see e.g. Rowland (2019: 145), and Suikkanen (2014: 86-89). And some, such as Bloom (2013), argue that babies make judgments of good and bad when they are as young as 3 months old.

[61] We might think that once we have any form of category we have correctness—see e.g. Finlay (2010: 332)—and since gender categories are categories, once we learn how to use them we have learnt how (we think) to apply them more or less correctly.

Furthermore, I believe that the view that cis people, in addition to trans and non-binary people, understand theirs and others' genders in terms of fit and correctness is both plausible and explains what cis people think and say. For instance, it seems that if one is a girl or woman (cis or not) one will think that it is correct, fitting, or appropriate (or at least not unfitting, incorrect, or inappropriate) to categorize one as a girl or woman rather than as a boy or man; for instance, that it is fitting to refer to one with she/her pronouns and to group one with the women or girls rather than the boys or men when there is such a grouping into genders, such as when the boys are asked to line up together and the girls are asked to line up together in school.[62] Serano (2016: 87–88) reports that when she asks public audiences whether they would change their gender for ten million dollars almost everyone says that they would not. As I explained in §2, when we judge that particular actions are not fitting in particular ways, we are often willing to bear large costs not to take those actions and we often have extremely negative experiences if we do perform these actions. To use the example that I introduced in §2, suppose your boss at the Amazon Warehouse will punish you if you do not praise Jeff Bezos. If you believe Bezos is anything but praiseworthy and you praise him to avoid punishment, you may feel ashamed, that you have been inauthentic or betrayed your values. And your view of Bezos may lead you to refrain from praising him even when refraining from praising him would lead to significant costs. So, a good explanation of why most people would not change their gender for ten million dollars is that (in part) they judge that if they changed their gender, their gender would not fit them and/or that it would be unfitting for them to navigate the world as a different gender and for them to be treated as a different gender by others. Serano (*ibid*: 88) says that when she asks people *why* they would not change their gender for ten million dollars, they respond by saying something like '"It just wouldn't be *right*"'. This is exactly what we would expect people to say if they thought that their taking this deal would involve their doing something, or lead to their doing something, unfitting. For instance, if our Amazon factory worker who doesn't think that it is fitting to praise Bezos (since he is not praiseworthy) was asked why they chose to refrain from praising Bezos at the cost of having their pay significantly docked, they might similarly say that 'it just wouldn't be right to praise him'.

7 | CONCLUSION

In this paper I've proposed a new descriptive or hermeneutical account of our gender judgments and gender concepts. This fitting treatment account (FTA) understands gender thought and talk as normative but not as morally, politically, prudentially, or all-things-considered normative thought and talk; rather, it understands gender thought and talk as thought and talk about the normative matter of the gender it is fitting to treat someone as. In §2 I argued that FTA explains why we encounter seemingly intractable disagreements about gender that outstrip descriptive agreement and why it is coherent to reject definitions of genders in terms of any set of descriptive non-normative features. In §3 and §4 I argued that competing non-normative contextualist and polysemic accounts of gender do not have these advantages and in fact struggle to explain many genuine disagreements about gender, and that FTA in fact has all the virtues of these views. In §4-5 I argued that competing normative contextualist and expressivist accounts of gender encounter problems, they make it impossible for us to judge that someone is not gender G but to judge that we should treat them as a G, or to in fact treat them as a G for other (moral or political) reasons, and this renders these views implausible. These problems derive from the fact that normative

[62] See *supra* note 57.

contextualist views put the wrong type of normativity into our gender judgments and because we can plan to treat someone as a G even when we take it to not be fitting or correct to do so for purely moral reasons. Other problems for normative contextualism have been noted elsewhere[63] but other problems for expressivism have not been discussed. So, in §5 I argued that we have further reason to accept FTA rather than expressivism because the latter encounters the Frege-Geach problem in a way that FTA does not, expressivism struggles to explain the harms in misgendering as well as FTA, and although FTA explains why many trans and non-binary people understand and explain their genders in terms of fit, expressivism does not. Finally, in §6 I showed that FTA does not seem to face any other serious problems. If my arguments are sound, I have shown that FTA merits serious consideration as an illuminating descriptive account of our current gender judgments, concepts, and our gender thought and talk in general.*[64]

REFERENCES

Ashley, F., & Ells, C. (2018). In favour of covering ethically important cosmetic surgeries: facial feminisation surgery for transgender people. *American Journal of Bioethics*, *18*, 23–25. https://doi.org/10.1080/15265161.2018.1531162

Ásta. (2018). *The Categories We Live By*. Oxford: Oxford University Press.

Bach, K. (2018). Loaded Words: On the Semantics and Pragmatics of Slurs. In D. Sosa (Ed.), *Bad Words*. Oxford: Oxford University Press.

Bach, T. (2012). Gender Is a Natural Kind with a Historical Essence. *Ethics*, *122*, 231–272. https://doi.org/10.1086/663232

Badgley, J., & Other Contributors. (2021). The Gender Dysphoria Bible. https://genderdysphoria.fyi/gdb

Barnes, E. (2020). Gender and Gender Terms. *Noûs*, *54*, 704–730. https://doi.org/10.1111/nous.12279

Barnes, E. (forthcoming). Gender without Gender Identity: The Case of Cognitive Disability. *Forthcoming in Mind.*, https://doi.org/10.1093/mind/fzab086

Bettcher, T. (2013). Trans Women and the Meaning of 'Woman. In A. Soble, N. Power & R. Halwani (Eds.), *Philosophy of Sex: Contemporary Readings*, 6th edition (pp. 233–250). Rowman & Littlefield.

Bettcher, T. (2014). Trapped in the Wrong Theory: Rethinking Trans Oppression and Resistance. *Signs: Journal of Women in Culture and Society*, *39*, 383–406. https://doi.org/10.1086/673088

Bloom, P. (2013). *Just Babies*. Crown.

Brentano, F. (1969) [1889].*The Origin of our Knowledge of Right and Wrong*. London: Routledge.

Broad, C. D. (1930/1956). *Five Types of Ethical Theory*. London: Routledge and Kegan Paul.

Butler, H. (1990). *Gender Trouble*. New York: Routledge.

Camp, E. (2013). Slurring Perspectives. *Analytic Philosophy*, *54*, 330–349. https://doi.org/10.1111/phib.12022

Combs, CK. (2019) What am I? In Rajunov and Duane (2019a).

Cull, M. (2019). Against Abolition. *Feminist Philosophy Quarterly*, *5*, 1–16. https://doi.org/10.5206/fpq/2019.3.5898

Dembroff, R. (2018). Real Talk on the Metaphysics of Gender. *Philosophical Topics*, *46*, 21–50. https://doi.org/10.5840/philtopics201846212

Dembroff, R. (2021). Escaping the Natural Attitude About Gender. *Philosophical Studies*, *178*, 983–1003. https://doi.org/10.1007/s11098-020-01468-1

Dembroff, R., & Wodak, D. (2018). He/She/They/Ze. *Ergo*, *5*, 371–406. https://doi.org/10.3998/ergo.12405314.0005.014

Diaz-Leon, E. (2016). Woman as a Politically Significant Term: A Solution to the Puzzle. *Hypatia*, *31*, 245–258. https://doi.org/10.1111/hypa.12234

[63] See *supra* note 42.

[64] I would like to thank audiences at the University of Leeds, the Chapel Hill Normativity Workshop, the Frankfurt Metaethics Conference, Social Ontology 2021, and the SWIP Session at the Joint Session of the Aristotelian Society and Mind Association for extremely helpful comments on previous versions of this paper. I would also like to thank Graham Bex-Priestley, Jessica Isserow, Nick Laskowsi, Ashley Shaw, Pekka Väyrynen, and an anonymous referee for amazingly useful substantive comments on previous drafts of this paper.

Erlick, E. (2019). Rethinking Non/Binary. In Rajunov and Duane (2019a).
Evers, D., & Streumer, B. (2016). Are the Moral Fixed Points Conceptual Truths? *Journal of Ethics and Social Philosophy*, *1*, 1–9. https://doi.org/10.26556/jesp.v10i1.183
Finlay, S. (2010). Recent Work on Normativity. *Analysis Reviews*, *70*, 331–346. https://doi.org/10.1093/analys/anq002
Geach, P. T. (1965). Assertion. *The Philosophical Review*, 74, 4, 449–465.
Gregory, A. (2021). *Desire as Belief*. Oxford: Oxford University Press.
Haslanger, S. (2012). *Resisting Reality*. Oxford: Oxford University Press.
Hom, C. (2010). Pejoratives. *Philosophy Compass*, *5*, 164–185. https://doi.org/10.1111/j.1747-9991.2009.00274.x
Howard, C. (2018). Fittingness. *Philosophy Compass*, *13*, 1–14. https://doi.org/10.1111/phc3.12542
Howard, C. (2019). The Fundamentality of Fit. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics Volume 14*. Oxford: Oxford University Press.
Howard, C., & Rowland, R.A. (2022). Fittingness: A User's Guide. In C. Howard and R. A. Rowland (Eds.), *Fittingness*. Oxford: Oxford University Press.
Iantaffi, A., & M. J. Barker (2018). *How to Understand Your Gender: A Practical Guide for Exploring Who You Are*. London: Jessica Kingsley.
Jenkins, K. (2016). Amelioration and Inclusion: Gender Identity and the Concept of Woman. *Ethics*, *126*, 394–421. https://doi.org/10.1086/683535
Jenkins, K. (2018). Toward an Account of Gender Identity. *Ergo 5*, *27*, 713–744. https://doi.org/10.3998/ergo.12405314.0005.027
Jacobson, D. (2011). Fitting Attitude Theories of Value. *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition). E. N. Zalta (Ed.), URL =<https://plato.stanford.edu/archives/spr2011/entries/fitting-attitude-theories/>.
Jeshion, R. (2018). Slurs, Dehumanization, and the Expression of Contempt. In D. Sosa (Ed.), *Bad Words*. Oxford: Oxford University Press.
Kapusta, S. (2016). Misgendering and its Moral Contestability. *Hypatia 3*, *3*, 512–519. https://doi.org/10.1111/hypa.12259
Laskowski, N. G. (2020). Moral Constraints on Gender Concepts. *Ethical Theory and Moral Practice*, *23*, 39–51. https://doi.org/10.1007/s10677-020-10060-9
McGrath, S. (2021). The Metaethics of Gender. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics Volume 16*. Oxford: Oxford University Press.
McKitrick, J. (2015). A Dispositional Account of Gender. *Philosophical Studies*, *172*, 2575–2589. 10.1007/s11098-014-0425-6
Mikkola, M. (2019). Feminist Perspectives on Sex and Gender. *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), E. N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2019/entries/feminism-gender/>.
Miller, A. (2003). *An Introduction to Contemporary Metaethics*. Cambridge: Polity.
Peters, T. (2021). *Detranisition, Baby*. New York: Penguin.
Rabinowicz, W., & T. Rønnow-Rasmussen (2004). The Strike of the Demon: On Fitting Attitudes and Vald'. *Ethics*, 114, 3, 391–423. https://doi.org/10.1086/381694
Ross, W. D. (1939/1949). *Foundations of ethics*. Oxford: Clarendon Press.
Rowland, R. A. (2019). *The Normative and the Evaluative: The Buck-Passing Account of Value*. Oxford: Oxford University Press.
Rowland, R. A. (2020). *Moral Disagreement*. Abingdon: Routledge.
Saul, J. (2012). Politically Significant Terms and Philosophy of Language: Methodological Issues. In S. Crasnow and A. Superson (Eds.), *Out from the Shadows*. Oxford: Oxford University Press.
Schroeder, M. (2008). *Being For*. Oxford: Oxford University Press.
Schroeder, M. (2010). *Noncognitivism in Ethics*. Abingdon: Routledge.
Serano, J. (2016). *Whipping Girl*. Seal press.
Sexsmith, S. (2019). Combing Out as Your Nibling. In Rajunov and Duane (2019a).
Shaw, A. (forthcoming). Desire and What It's Rational to Do. Forthcoming in Australasian Journal of Philosophy.
Skorupski, J. (2010). *The Domain of Reasons*. Oxford: Oxford University Press.
Stryker, S. (2008). *Transgender History*. Seal Press.
Suikkanen, J. (2009). Buck-passing accounts of value. *Philosophy Compass*, *4*, 768–779. https://doi.org/10.1111/j.1747-9991.2009.00253.x

Suikkanen, J. (2014). *This is Ethics*. Oxford: Blackwell.
Turner, J. (2021). Jess Phillips: "The only way a woman will become Labour leader is it men don't stand". *The Times*, July 17, 2021.
Violet, M. (2018). *Yes, You Are Trans Enough: My Transition from Self-Loathing to Self-Love*. London: Jessica Kingsley.
Way, J. (2013). Value and Reasons to Favour. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics Volume 8*. Oxford: Oxford University Press.
Weiss, S. (2018). 9 Things People Get Wrong about Being Non-Binary. *Teen Vogue*, February 15, 2018. https://www.teenvogue.com/story/9-things-people-get-wrong-about-being-non-binary
Williams, J. R. (2016). Vagueness as Indecision. *Aristotelian Society Supplementary Volume XC*, 285–309. https://doi.org/10.1093/arisup/akw001
Woods, J. (2014). Expressivism and Moore's Paradox. *Philosophers' Imprint*, *14*, 1–12.

How to cite this article: Cosker-Rowland, R. (2024). The normativity of gender. *Noûs*, *58,* 244–270. https://doi.org/10.1111/nous.12453