THE ETHICS OF GENDERED NAMING, THE MORAL DANGER OF LOVE AVOWALS, AND THE MORAL VALUE OF INSULTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Adriene Takaoka

August 2023

© 2023 Adriene Takaoka

THE ETHICS OF GENDERED NAMING, THE MORAL DANGER OF LOVE AVOWALS, AND THE MORAL VALUE OF INSULTS

Adriene Takaoka, Ph. D.

Cornell University 2023

I pragmatically and morally analyze three categories of ubiquitous speech: gendered naming, love avowals, and insults. In Chapter 1, I argue that our naming practices are out of step with widely held social values about gender equality. I identify what I call the BioSex Logic of Naming, according to which strong social norms prescribe that first names track presumed biological sex. I argue gender-specific names and gender-specific pronouns are natural bedfellows, both of which play a crucial, companionate role in reinforcing gender essentialism and maintaining gender hierarchy. I conclude that we have strong *pro tanto* reasons to jettison both gender-specific pronouns and gender-specific names while also protecting binary trans* people who are at beholden to the current system for gendered uptake. In Chapter 2, I argue love avowals have a morally dark side. At their best, love avowals, such as, "I love you" are welcome expressions of affection and devotion. Love's positive public image makes love avowals an appealing tool of manipulation. But, like the Trojan Horse, its surface appearance conceals ugly realities. This simplistic veneer diverts our attention away from the normative profundity of love avowals, which often function pragmatically as solicitations, illegitimate demands, excuses, exonerations, and the like. Love avowals conceal tacit bids for commitments that one may not entertain if they were communicated explicitly. I illuminate the pragmatic function of love avowals and argue we ought to be more critical of them, as they play a key role in consequential social scripts, some of which facilitate abuse. In Chapter 3, I argue, contrary to dominant thought, insults are not inherently morally wrong. Insults can

have a unique organizing role in social justice movements and reimagining social hierarchies of race and gender. Extant theories generalize the nature of insults from a narrow subset of insults and conceive of them as invariably problematic. This theoretical commitment to the immorality of insults precludes the possibility of developing an ethics of insults. I develop a morally neutral theory of insults that identifies an overlooked variety, third-personal insults, and allows a distinction between permissible and impermissible insults.

BIOGRAPHICAL SKETCH

Adriene Takaoka completed their undergraduate work in Philosophy at Wellesley College, graduating with honors in 2015. They received their master's degree in philosophy from Cornell University in 2018.

ACKNOWLEDGMENTS

I extend my gratitude to the Sage School of Philosophy for both their intellectual and financial support throughout my graduate career, especially for The Dissertation Completion Fellowship and for supporting my summer conference travel every year. I'm indebted to the Office of Inclusion and Student Engagement at Cornell University for awarding me The Dean's Excellence Fellowship, which funded the majority of my graduate career. I'm also indebted the Cornell University Graduate School for awarding me The Provost Diversity Fellowship, which relieved me of my teaching duties for a semester, allowing me to focus on my own research. I'm grateful to my dissertation committee: Kate Manne, W. Starr, Julia Markovits, Mary Kate McGowan, and Sarah Murray. Special thanks to my co-chairs, W. Starr and Kate Manne, whose unwavering support and confidence in my philosophical ability and my outside-the-box philosophical orientation and propelled me through difficult times when I felt at sea. Thank you for modeling intellectual independence and humility and for supporting mine. Working with you was a privilege.

TABLE OF CONTENTS

Chapter 1
The Ethics of Gendered Naming: Naming Yourself and Naming Others

1. Introduction

Many trans*[1] people are unhappy with their given first name, because either it *misgenders* them or because it *genders* them simpliciter. Choosing a new name can be morally daunting. This is because current naming conventions adhere to what I call the BioSex Logic of Names, according to which strong social norms prescribe that first names signal the biological sex one was assigned at birth. Babies with penises get boy names and babies with vaginas get girl names. Intersex babies who have both a penis and a vagina, or ambiguous genitalia, are assigned a sex a birth, and subject to binary naming norms. To appreciate the moral bind confronting many trans people, consider the following scenario. X is a trans* man whose given name is "Amanda." You might think X should simply change his legal name to a recognizably male name, like "Joshua." On the one hand, X may desire exactly that. Given that names are assumed to track one's genital status, having an unequivocally male name would help X both to avoid being misgendered, and to avoid being forced to disclose that he is transgender (through a discrepancy between his name and appearance). Yet X may be reluctant to choose a binary name because he rejects the underlying essentialist ideology of BioSex norms, which relegate people to false binary gender categories. By assuming the binary name "Josh," X would be complicit in the very system that caused his bind in the first place and that will continue to harm other trans* people. This is a moral dilemma: X desires social uptake as a man—and names significantly help obtain uptake—but he rejects that names should track assigned biological sex.

Although this issue is particularly salient for trans* people, the moral ramifications of BioSex naming norms are germane to all of us for at least two reasons. First, those of us who are parents are usually tasked with naming our children. We are the very people who give names that turn out to be

[1] This is an umbrella term that includes transgender, agender, and bi-gender people. See Stryker, *Transgender History*, 11.

unwelcome and even harmful. Presumably we should aim to name our children in ways that will not come to aggrieve them or contribute to their oppression. After all, there are reasons we don't typically give our children names like "Coca Cola." We don't want to them to have to face the scrutiny of their peers, bias, or shame. Second, as I shall argue, we ought to become skeptical about BioSex naming norms, because of their negative consequences for both trans* and cisgender people. Gender-specific names tacitly communicate that biological sex is binary and essential to who we are fundamentally. Note that I am not claiming that ascriptions of gender are never relevant; rather, gender-specific names communicate that gender is relevant far more often than it is, thereby tacitly endorsing an untenable essentialist view about gender. This is true independently of your metaphysics of gender; whatever you might believe makes it the case that someone has a gender, it's not at all clear that first names should communicate gender information.[2]

Most work on the ethics of gender ascriptions has focused on pronouns, which is understandable given their recent spotlight in political discourse. Interestingly, very little has been said in feminist philosophy of language about gender-specific naming, despite its ubiquity and its being notably out of step with widely held values about gender equality. I'll argue that the practice of gender-specific naming is harmful to both trans* and cisgender people—to different extents—because it perpetuates a kind of gender essentialism that undermines agency and wellbeing, and we therefore have a *pro tanto* reason to use gender-neutral names. The paper will unfold as follows. First, I'll motivate the BioSex Logic of Naming. Second, I'll use Dembroff and Wodak's insightful work on gender-specific pronouns as an inroad to thinking about gender-specific names, because their considerations illuminate much about what is wrong with gender-specific names. Third, I'll argue, against Dembroff and Wodak, that not only does their analysis extend to names, but gender-specific

[2] I'm assuming strict gender essentialism isn't a defensible metaphysics of gender because such as view is demonstrably false.

pronouns and gender-specific names are fundamentally intertwined and mutually reinforcing. To vivify this, I extend Bettcher's work on transphobic violence, arguing that gender-specific names constitute a kind of *linguistic attire* that plays a crucial role in the communicative relation between appearance (attire) and reality (sexed bodies) that gives rise to transphobic violence. I then evaluate strategies and complications for implementing a gender-neutral naming practice. I go on to consider the ethical upshots of the view for trans people who find themselves in the moral dilemma of desiring the gendered uptake to which they are entitled while simultaneously rejecting the gender essentialism underlying the practice of BioSex naming. These are genuinely competing moral values from which we cannot simply derive a universal moral rule about what to do. Rather, I aim to assist trans* people in making clear-headed, empowered choices under unideal circumstances by highlighting the choice architecture confronting them. Finally, I'll urge cisgender people to take responsibility for changing this harmful practice, in coalition with trans* people.

2. The BioSex Logic of Names

The claim that our naming practice is problematically gendered may not be intuitive, so I'll motivate the view here at the outset. We expect presumed biological boys and men to have *male names* and presumed biological girls and women to have *female names*. This is a binary system, so intersex people are excluded from this paradigm, though they're nevertheless subject to the general rule that names indicate genital status. A quick search will show that the U.S. Census divides data about names into binary gender categories. Likewise, a quick search for baby names almost always yields lists of names according to gender category. In an analysis of U.S. data, Leiberson et al. found that 97% of "female-appropriate" names are given to children assigned female at birth, and the converse is also true: 97% of "male-appropriate" names are given only to children assigned male at birth.[3] The concordant relationship between names and gender has been noticed by those who

[3] Leiberson, Dumais, and Baumann, "The Instability of Androgynous Names," 1261.

collect data about online subjects. Indeed, the most reliable way to infer the gender of online subjects is through analysis of first names.[4] Furthermore, first names and pronouns are natural bedfellows: we infer pronouns from names, and we infer what kind of name a person has (gender-specific) from their pronouns. In college classroom settings, for example, students often read the work of authors about whose gender presentation they're ignorant. Nevertheless, they confidently infer pronouns from the first names of authors.

You might think that gender-neutral names, apparent deviations from the norm, tell against this picture. But actually, they support it. Gender-neutral names tend to be unidirectional. Girls can have boys' names, but boys cannot have girls' names. You might know, as I do, a cisgender woman with a conventionally male name like "Michael," but you'll probably never meet a cisgender man named "Nancy" or "Veronica," which shows that traditionally female names are untenable as gender-neutral names for boys. In turn, the untenability of traditionally girls' names as gender-neutral names for boys, suggests that girls' names are associated with undesirable characteristics associated with stereotypes about biological females.[5] Empirical evidence demonstrates this bias. In a 2012 randomized double-blind study, science faculty from research-intensive universities rated the application materials of a student, who was randomly assigned either a male or female name, for a laboratory manager position. Faculty rated the male applicant as significantly more competent and hirable than the otherwise identical female applicant. They also selected a higher starting salary and offered more career mentoring to the male applicant. Furthermore, in their study of androgenous names, Lieberson et al. found that parents are more inclined to use androgynous names for daughters than for sons, which causes androgynous names to be unstable. They call this the

[4] Admittedly, the first study conflates sex with gender presentation by confirming the sex of the subject based on their profile picture, though this doesn't obviously undermine the expected congruity between name and gender appearance. Liu and Ruths, "What's in a Name?"; Bérubé et al., "Wiki-gendersort: Automatic Gender Detection Using First Names in Wikipedia."

[5] Moss-Racusin et al., "Science Faculty's Subtle Gender Biases."

"contamination" problem whereby "the advantaged have a greater incentive to avoid having their status confused with the disadvantaged." If an androgynous name is given to daughters with regularity, then parents will avoid using that name for sons.[6] Given this, we can conclude that many gender-neutral names are neutral in the sense that "he" is gender-neutral—not at all.[7] The contamination problem shows that gender-specific names are a key tool for maintaining gender hierarchy.

Finally, the normativity of BioSex naming is evinced by our reactions to discrepancies between names and apparent genital status. Suppose there are two children named Sam in an elementary school classroom. Suppose further that Sam isn't shorthand for either student; Sam is the full first name of both—not Samantha and Samuel. When the need arises to distinguish them, we'll reach for "girl Sam" and "boy Sam" rather than, say, "tall Sam" and "short Sam," "Sam R." and "Sam J.," or even "Black Sam" and "White Sam." When names fail to communicate gender, our first instinct is to distinguish them by filling in the gender information that's missing in the name. Relatedly, in an interview with the sociologist R. Connell, a trans woman named Julie reported her strange experience at her telephone-based customer service job. Customers would make sense of the discrepancy between her voice and her name by mishearing her name and substituting it for a male-sounding name.[8] This illustrates the strength of our normative expectations.

Some deviations elicit mere surprise while other instances evoke shock and disgust. For example, imagine a feminine-presenting cisgender girl named "Jordan." Her name will sometimes evoke surprise, which indicates a deviation from the rule. However, so long as Jordan looks sufficiently "like a girl" such that her gender presentation matches her presumed genital status, then her deviant name won't incite further curiosity or suspicion. Because there's no apparent asymmetry

[6] Leiberson, Dumais, and Baumann, "The Instability of Androgynous Names," 1285.
[7] Dennis Baren, *What's Your Pronoun?,* 39-78.
[8] Connell, "Doing, Undoing, or Redoing Gender," 41-42.

between Jordan's gender presentation and her presumed genital status, her aberrant name will be, for example, attributed to a choice made by her quirky parents or perhaps regarded as a cute idiosyncrasy. Meanwhile, the mere counterfactual world in which a masculine-presenting person (who is assumed to have a penis) walks into a room and says, "Hey, I'm Amanda" would be sure to evoke shock, ridicule and even disgust.

3. Gender-Specific Pronouns

My thesis that the practice of gender-specific naming is harmful and therefore ought to be jettisoned may seem radical. But I think it is actually rather vanilla, in the sense that it follows from basic values about gender equity and fairness. To begin to carve out my view, it will be instructive to consider literature on the ethics of gender-specific pronoun use, because those considerations reveal many of the moral problems with gender-specific names. In their article, "How Much Gender is Too Much Gender," Dembroff and Wodak (D&W) aim to show that the contemporary use of gender-specific pronouns, honorifics, suffixes, and generic terms are not as innocuous as it might appear. To show this, they ask us to turn to a fictional language, "Renglish," created by Douglas Hofstadter, in which pronouns are not gender-specific, but race-specific.

> Hofstadter imagines a world with race-specific English terms (hereon 'Renglish') that corresponds to the actual world's gender-specific English terms. Writing in the voice of "William Satire"—a fictional person whose name, prose, and arguments resemble those of the columnist William Safre—Hofstadter considers the 'negrists' who would eliminate certain race-specific terms and thereby "radically change our language in order to 'liberate' us poor dupes from its supposed racist bias". These terms fall into four categories, which have gender-specific counterparts in English.

	Renglish	*English*
Pronouns	*whe*; *ble*	*he*; *she*
Honorifics	*Master*; *Niss*	*Mister*; *Miss*
Suffixes	*-oon*	*-ess*
Generic terms	*whiteslaughter*	*manslaughter*

Hofstadter's analogy elicits the intuition that English should be no more gender-

> specific than it is currently race-specific with respect to these parts of speech. That is, just as we should not have these distinct race-specific pronouns ("whe"), generic terms ("chairwhite"), honorifics ("Nrs."), and suffixes ("oon") in English, we should not have their gender-specific equivalents.[9]

D&W highlight three reasons in support of the conclusion that Renglish is morally indefensible. First, it's exclusionary and distorting insofar as it treats racial categories as binary, exhaustive, and mutually exclusive. For example, Asian and Native people do not fit into this picture, and it's not at all clear which pronouns would be used for people of mixed race. Likewise, gender-specific pronouns exclude people who are intersex, non-binary, agender, bi-gender, etc. D&W hold that it's infeasible to rectify this by simply proliferating either race-specific or gender-specific pronouns to include every race or gender identity that is left out or distorted by pronouns. This is because pronouns are what linguists call a "closed class," or cognitively primitive speech, which makes learning new pronouns prohibitively difficult. This linguistic confusion would increase the risk of making mistakes about others' pronouns, and we should seek to minimize misgendering.[10]

Second, Renglish generates a deceive or disclose bind for speakers, whereby speakers are often forced by grammar to either disclose information about race, or to deceive their interlocutors by deploying circumlocutions that avoid pronouns. For example, "Imagine you live in a society that speaks Renglish. One day at work, you are making small talk with a colleague when they ask you whether you have any summer vacation plans. You say you have a vacation planned with your partner. Your colleague responds using the wrong race-specific pronouns for your partner: they didn't know that you're in a mixed-race relationship. Perhaps you'd prefer to keep this private—you think it's no one else's business, and you'd rather not have to engage with those who think that

[9] Dembroff and Wodak, "How Much Gender is Too Much Gender?," 363-364.
[10] Ibid, 366-367.

mixed-race relationships are morally wrong or disgusting."[11] In this case, you are forced by grammar to either disclose information about race that you're entitled to keep private or to deceive your interlocutor by letting them go on thinking that you have a partner of the same race. This is also the case with gender-specific pronouns. The same bind can be illustrated by simply swapping out a gender-specific pronoun for the race-specific pronoun. Grammatical norms, D&W hold, should not violate our privacy and autonomy, and therefore both race-based and gender-based pronouns are structurally morally problematic.[12]

Third, Renglish, stigmatizes and stereotypes people: "Renglish stigmatizes blacks in part by making 'white' the generic term for all persons in phrases like 'chairwhite' and sentences like 'All whites are created equal': making 'white' the default treats non-whites as different, defective instances of persons. The same concern carries over to generic uses of 'man' for persons; feminist philosophers have long objected to generic uses of 'he' and 'man' in English."[13] Building racial identities into the very grammar of the language communicates that race is always relevant to explaining everything one does, from doing dishes to going to the moon. The authors suggest that these problematic implications are generated by Grice's Maxim of Relation, whereby a speaker's contribution to the conversation is expected to be relevant to the conversation at hand.[14] Empirical evidence on the relation between social cognition and the effect of language bolsters their view. For example, one study showed that the use of gender-neutral pronouns "reduces the mental salience of males," which in turn is associated with more positive public attributes towards females and

[11] Ibid, 367.
[12] Ibid, 367-368
[13] Ibid, 365.
[14] Dembroff and Wodak point readers to further alternative mechanisms that conveys social meaning regardless of speaker intention. Dembroff and Wodak, "How Much Gender," footnote 24. I would add that Mary Kate McGowan's work on exercitive speech is also a good candidate for explaining how pronouns (and names) do this work. She argues speech constitutes the harm of oppression when (a) it enacts an oppressive social norm, (b) the norm is followed and (c) harm results from the following of that norm. Using a pronoun would count as a covert exercitive that enacts oppressive norms. McGowan, *Just Words,* 82-123.

LGBTQ persons, and another study showed that an object's (human and inanimate) grammatical gender categorization impacts cognition by influencing "conceptual gender" expectations for the object. Another series of studies showed that grammatical gender systems bias speakers' memories, descriptions, and assessments.[15]

While the authors are careful not to conflate race and gender, acknowledging their differential genealogies and social meanings, they argue that this specific analogy between race-based and gender-based pronouns is apt, for exactly the reasons they specify: stigma, stereotypes, exclusion, privacy, etc. Thus, they conclude, we have good reason not to use any gender-specific pronouns to refer to anyone, regardless of their gender identity.

In earlier work, D&W make a stronger version of the same argument: we have a negative *moral duty* not to refer to anyone using gender-specific pronouns, and the best way to do so (given the prudential reason against proliferating pronouns that gives rise to the moral problem of misgendering), is to use "they" universally. This duty is rooted in the idea that we don't have a positive duty to affirm gender. After all, we don't have a duty to affirm other important aspects of people's identities, such as race and occupation, through pronoun use. Rather, we must not *deny* peoples' gender identities. The duty not to misgender is rooted in considerations about respect, access to resources, intelligibility, and gender essentialist ideology. Disrespect via misgendering implies we may withhold from trans* people resources to which they are entitled, and misgendering thwarts intelligibility by subjecting trans* people's conduct to interpretive guides, or blueprints, that they justifiably reject.[16] Using the egalitarian "they" allows us to avoid gendering (and therefore misgendering) *anyone*, and in doing so we avoid endorsing essentialist gender ideology in the way gender-specific pronouns do.

---

[15]The first two examples are in footnote 22 of "How Much Gender is Too Much Gender" and the second is from Dembroff and Wodak's earlier article, "He/She/They/Ze," 396.
[16] Dembroff and Wodak, "He/She/They/Ze," 378-379.

The reasoning in D&W's earlier and later work overlaps significantly, but their commitment to "duty" in their earlier work renders their argument much stronger than their later work and is certainly stronger than the one I will make here. Their more recent work, "How Much Gender is too Much Gender?" does not mention moral duties at all, so I interpret their view to be that we have *pro tanto* moral reasons to not use gender-specific pronouns for anyone and not that we have a *moral duty* to refrain from doing so such that using a gender-specific pronoun would count as *universally morally wrong*. Nevertheless, the upshot of their earlier, stronger view in "He/She/They/Ze"—that we should all use the gender-neutral pronoun "they"—does not rely on us having moral duties, and is therefore perfectly compatible with their later view that I rely on here. Therefore, I take the upshot from their earlier work and the race-analogy argument to form a single coherent view: We shouldn't use gender-specific pronouns for anyone; we should use "they" instead.[17]

4. From Gender-Specific Pronouns to Gender-Specific Names

I'll now argue that Dembroff and Wodak's view strongly suggests that we should also be worried about the negative moral ramifications of gender-specific names, because they do all the same things that gender-specific pronouns do. To appreciate this, suppose D&W's view was implemented, and we successfully ceased to use any gender-specific pronouns to refer to anyone regardless of gender identity and, instead, we use the pronoun "they" universally. Now imagine a scenario in which you tell your boss you're going on vacation with your partner. Your boss might respond, "Cool, what's your partner's name?" If you're a gay woman and you don't want to disclose that, then you're in a bind that pronouns can't solve: Their name is Nancy. You can only retain your privacy by deceiving your boss. If deception doesn't seem feasible to you, then you must disclose,

[17] In their earlier paper "He/She/They/Ze," Dembroff and Wodak cite considerations of gender essentialist ideology with its attendant stigma and stereotyping as a reason against the use gender-based pronouns, whereas, in their later paper, "How Much Gender is Too Much Gender?," they drop any discussion of ideology and appeal only to stigma and stereotyping considerations. Presumably, they can cash out all their reasoning empirically without introducing complications that come with the ideology literature.

and we are back in the deceive or disclose bind that D&W want to remedy.

Interestingly, D&W deny that their view extends to gender-specific naming practices.

> In a sense, gender-specific proper names might seem to rid someone of their privacy and autonomy. For instance, say you are communicating online, and have to disclose your proper name. If your name is "Rachel', this means implicitly that you are a woman. And since you may have good reason not to disclose your gender identity (without deception), this can be problematic. But notice that you can legally change your proper name, and you can also adopt a gender-neutral nick name (like 'Ray') without even filing paperwork. Unlike with pronouns, honorifics, adjectives, and so on, generally others do not simply assume one's proper name based on one's appearance, posture, or voice. We are authorities over our proper names: as a result, we have much more control over what this name will be. So perhaps there is some reason for parents to use gender-neutral names for their children, but it's not decisive. Moreover, telling parents to pick a gender-neutral name can sometimes veer close to the objectionable demand that they *anglicize* their child's name. Consider Iceland, where most individuals can only have a patronym whereby the suffix '-son' or '-dóttir' (depending on the declared sex of the child) is added to the genitive form of the father's name. It's one thing for Icelanders to argue for Iceland to allow individuals to adopt gender-neutral hereditary surnames. But it's another thing entirely for English speakers to demand that Icelandic immigrants do the same.[18]

The authors hold that their considerations about privacy and autonomy do not extend to gender-specific proper names for three reasons. First, they argue we are the authorities over our proper names insofar as we can legally change our proper names or simply adopt nicknames. I take this to mean that pronouns are a closed class of speech and therefore unamenable to change, whereas proper names are an open class that admits of modification. Therefore, the negative normative ramifications of pronouns do not apply to names. Second, others don't simply assume our proper name based on our appearance, posture, or voice. Third, such a view involves problematic moralizing, such as asking parents to *anglicize* their child's name or demanding Icelandic immigrants refrain from using gender-specific names. All these justifications are mistaken.

First, the authors are overly optimistic about the extent of control we wield over our first names and this is empirically demonstrably. Legally, changing your name in the U.S. is daunting. It

[18] Dembroff and Wodak, "How Much Gender is Too Much Gender?," 372.

requires extensive documentation, and bureaucratic fees, which are not trivial barriers for many trans* people.[19] The non-profit think tank Movement Advancement Project reports that 68% of transgender and nonbinary adults lack any government-issued ID that matches their name and gender identity.[20] Additionally, in the U.S., you must be 18 years-old to change your legal name without parental consent. These obstacles can be detrimental for some people, like 17-year-old, Gwen Araujo, a woman who was murdered for being transgender in 2002. Before she died, she experienced persistent harassment at school and had difficulty finding employment because her appearance as a girl did not match her legal name on job applications.[21]

Moreover, even when one does change their legal name, it does not follow that others then use it correctly. Medical offices are notoriously bad at updating their systems to consistently reflect new names, and patients must deal with the subsequent confusion potentially indefinitely. Suppose you're a trans* person who does have the resources to legally change your name, but you're a regularly cited author. Deadnaming will be a lifelong inevitability. Furthermore, changing one's name, legally or not, is socially painstaking. It requires repeated disclosure. While social media has made it easier to disseminate information to many people at once, updating one's name does not involve a one-time announcement—it's an ongoing, long-term process. This gives rise to the very privacy issue that D&W take it to solve. When trans* people change their names, informing people of their name change is often tantamount to "coming out." Why else would "Bill" change their

[19] Roughly 30% of transgender people live in poverty—two times the rate of non-transgender people. In a national survey, 35% of transgender and nonbinary people who hadn't changed their legal name, as well as 32% of those who had not updated the gender marker on their IDs, cited cost was the main barrier. Compared with otherwise comparable cisgender men, transgender women have significantly lower employment rates, lower household incomes, higher rates of poverty, greater Medicaid use, and increased likelihood of food insecurity. Non-cisgender AFAB individuals have significantly lower employment rates than otherwise similar cisgender women. Both AMAB and AFAB black individuals have significantly worse economic outcomes than non-cisgender white individuals. Carpenter, Lee, and Nettuno. "Economic Outcomes for Transgender People," 280-304; Movement Advancement Project and Center for American Progress. *Paying an Unfair Price.*

[20] Movement Advancement Project and Center for American Progress. *"*Paying an Unfair Price."

[21] Bettcher, "Evil Deceivers and Make-Believers," 44.

name to "Nancy," or vice versa? Even though, phenomenologically, this process of repeatedly coming out feels like an invasion of privacy, notice that no one in particular is to blame. After all, people need to be informed of your name change in order to refer to you properly. This privacy issue is generated by the BioSex naming system itself, which codifies the binary communicative relation between names and assigned biological sex.

It's also overly optimistic to think we can avoid this problem by adopting a nickname. Not only will contexts that require legal names remain problematic, but we often do not have direct control over our nicknames. Nicknames are often given and not chosen. We can expect that young people especially, would have a difficult time simply declaring a nickname and having it respected. One can easily imagine a child's siblings or schoolmates simply refusing or even ridiculing them for even trying to adopt a new name. Moreover, institutions such as universities often lack the infrastructure to reliably accommodate nick names. For example, suppose one has adopted a nick name, but this nickname is not registered in the university system, either because there isn't one or the system doesn't work well. On the first day of class, a professor might take attendance, and if you want to be counted, you'll be forced to acknowledge your legal name. A nickname would do fine in the online context that D&W mention, but presumably very few of us, both children and adults, live our lives exclusively online, among strangers. In summary, while names are technically changeable, they are socially obstinate.

Perhaps a more damning objection to the idea that gender-specific names are not as pernicious as pronouns is that it doesn't take into account the effects of a lifetime of already having had a gender-specific name. Trying to undo the stigma and stereotyping that comes with repeated gendering via naming is tantamount to attempting to un-ring a bell, to use Mary Kate McGowan's helpful metaphor. In her work on oppressive speech, McGowan stresses the asymmetric pliability of conversational norms, the phenomenon whereby it is easier to introduce conversational norms than

it is to subsequently undo them.[22] If names presuppose binary gender norms, then every utterance of a gender-specific name introduces gender norms into the conversation. It's difficult to remove or block these gender-essentialist presuppositions, especially in the case of names, I hold. Suppose I see my friend on campus and say, "Hi, Nancy!" It would be exceedingly cumbersome, socially awkward, and impractical for either one of us to attempt to block the presuppositions of her name. Imagine, I said, "hi, Nancy. Oh, but don't get me wrong, in referring to you by your name, I'm not endorsing the gender essentialism tacit in naming practices. Certainly, I don't think you're necessarily heterosexual, or fundamentally caregiving, or naturally suited to motherhood!" Likewise, it would be untenable for Nancy to respond to my greeting, "Hi, Nancy" by saying "That's false! I mean, it's not false that my name is Nancy, but it's false that I am someone to whom problematic gender norms apply!" We're stuck with the essentialist presuppositions that undergird gender-specific naming, even if we don't believe or support them. The social effects of repeated gendering don't simply cease when one changes their name, legally or otherwise. That bell has rung.

Recall the empirical evidence cited by D&W that shows an object's grammatical gender categorization—inanimate or human—impacts cognition by influencing "conceptual gender" expectations about the object. It's unclear whether it's possible to recategorize an object (in the minds of either speakers or hearers) once that object or person has already been categorized repeatedly via a gender-specific name throughout their lifetime. Also recall the study that showed that gender neutral pronouns reduce the mental salience of males, which in turn was associated with more positive public attributes towards females and LGBTQ persons. While the study tested only pronouns, it is not a major logical leap to assume blocking assumptions about people by omitting their names would yield similar results.

Not only do names (and pronouns) socially categorize you in the minds of others, but they

[22] Mary Kate McGowan, "Oppressive Speech," 403.

also shape how you perceive yourself. Incessant gendering—through names, pronouns, and other social practices—impacts one's own self-conception. Normatively inflected names influence who we take ourselves to be. As a result, many of us don't initially realize we are not cisgender, and some of us never do, because we are told repeatedly who we are—either a boy or a girl. If you have been successfully gendered, it might not even occur to you to resist that gendering, never mind change your name.

This lacuna in imaginative resources to interpret one's experiences is what Miranda Fricker calls hermeneutic injustice, "the injustice of having some significant area of one's social experiences obscured from collective understanding owing to a structural identity prejudice in the collective hermeneutical resource."[23] In other words, one of the harms of having a marginalized identity (realized or not) is having your life experiences be, to a significant extent, inscrutable and vexing. Talia Mae Bettcher adroitly calls this trans* phenomenological experience the "WTF." The WTF serves as a counter argument to "pristine philosophy" according to which the goal of philosophy is to import probing questions into an otherwise pristine, uncomplicated social world. On the contrary, Bettcher says, "We trans people live an 'everyday' shot through with perplexity, shot through with WTF questions. We live in the WTF. We did not need philosophy to uncover its perplexity. It was already there."[24] In summary, gender-specific names and pronouns play a crucial, companionate role in lodging trans* people firmly in the WTF.

To further elucidate the ways in which gender-specific pronouns and names shape our self-conceptions and worldview, I'm going to delve into Kukla and Lance's (K&L's) work on gender

[23] Fricker, *Epistemic Injustice*, 155. This kind of injustice is hinged on a pre-existing identity prejudice, so it only applies to trans* people. While BioSex naming, and it's underpinning gender essentialism, doesn't constitute an injustice to cisgender people, in Fricker's sense, I hold that it does harm cisgender people by limiting their imaginations about who they are and who they could be, which can lead to a host of other more concrete harms, like staying in abuse relationships or following destructive beauty standards.
[24] Bettcher, "What is Trans Philosophy?," 649-652.

ascriptions. While K&L only discuss gender ascriptions as they relate to pronouns, I take gender-specific names to be a paradigmatic case of gender ascription. According to K&L, "gender ascriptions function to organize social space, and support or undermine people's autonomy, bodily agency, and self-determination within this social space, in important ways."[25] While gender ascriptions have declarative grammatical form, they don't simply describe antecedent reality. Rather, they function primarily to alter and order social space. For example, "Hi, Nancy!" appears to be a simple greeting, and it is, but it also functions to locate Nancy in social space, independently of the speaker's intention to do so. When one is recognized as having a gender (by oneself or by others), one is positioned in normative space insofar as the social recognition of gender inserts us into a "complex web of normative expectations and pressures." K&L:

> Notice that what gender we are taken as having inflects nearly every aspect of how we are expected and demanded to negotiate the social and material world. It shapes how we are supposed to hold our body and modulate our voice, what clothes we are supposed to wear, how we are supposed to manifest sexual attraction and attractiveness, where and how we pee, what hobbies and jobs we are supposed to have, who we compete against in sports events and which sports we take up in the first place, what our relationship is to our children, and so forth. Even fetuses, once recognized as 'boys' or girls', are expected to become babies from whom certain nursery and clothing colors and emotions and behaviors are appropriate. Such norms are modulated by race, age, ability, class, body shape, and more—there is not just a single set of norms for each gender, but rather a complex and often contradictory web or norms in which we are differently positioned—but these structures of social significance are inescapable.[26]

On this view, gender ascriptions have profound material consequences. Given these profound material consequences and the basic norms of self-determination and bodily agency, K&L think we ought to have defeasible control over our placement within social space and social norms. From this, they conclude that first-person gender ascriptions are almost always entitled, and second and third person ascriptions that contradict first-person ascriptions almost never are. In other words, it is

[25] Kukla and Lance, "Telling Gender," 1.
[26] Kukla and Lance, "Telling Gender: The Pragmatics of Gender Ascriptions," 5-6.

permissible to refer to someone in the second or third person with a gendered pronoun when that person licenses you to do so, though, K&L hold that we can often infer pronouns from gendered appearance when we're unaware of someone's pronouns.[27] Even though K&L recognize the negative consequences of gender ascriptions, they nevertheless hold that we ought not to eliminate them at this particular moment in time, because doing so would undermine some people: "gender self-ascriptions are powerful tools that many people use to gain access to a range of interactions, sense-making resources, possibilities for self-expression, social opportunities, and forms of recognition and permission."[28] While K&L don't elaborate this point, I take them to be rightly trying to avoid undermining binary transgender peoples' first-person claims to gender membership. This, of course, is important, but I think it's mistaken to treat that commitment and the longer-term goal of eliminating harmful gender ascriptions as mutually exclusive.

Characterizing the nature of gender ascriptions, as K&L have done, and thinking about how we should proceed given what we know about gender ascriptions are two separate issues. The authors are right not to issue a universal prescription that we all stop using gender ascriptions immediately, an all-too-common impulse among ethicists. However, I would add that we can attend to both what we (trans* and cisgender people) should do right now given our unideal circumstances (defer to first-person gender ascriptions), while we also move toward eventually eliminating problematic gender ascriptions. While K&L are amenable to a future time in which gender roles are antiquated, they do not call for the elimination of gender-specific pronouns because doing so would sacrifice some people's well-being for the sake of a utopian future.[29]

This is a good place to remember that trans* people are not a univocal group, and some of us, especially those of us who are non-binary and agender, are unhappy with gender-specific

[27] Ibid, 21.
[28] Ibid, 12.
[29] Ibid.

pronouns at the very same time that we respect our binary trans* peers' gender ascriptions. Moreover, notice that if gender ascriptions are as socially powerful (and therefore dangerous) as K&L have argued they are, then we're always collectively sacrificing some people. If we don't seek their careful, strategic elimination, then we will continue to sacrifice the children who are growing up in this oppressive gender system and the adults who are suffering now. Finally, while names and pronouns are currently central mechanisms for obtaining gendered uptake, they are not the only possible way to do so. If these linguistic tools were no longer central mechanisms for communicating gender *for anyone*, then they wouldn't constitute the serious problem for binary transgender people that they do now. Those people for whom gender is important could communicate their gender through other practices that don't involve *gendering everyone* through grammatical fiat. Bringing about such a world is admittedly idealistic, but not naïve.

In summary, I disagree in principle (but not in spirit), with Kukla and Lance's view about *what we should do given the nature of gender ascriptions*, but I think their *analysis of gender ascriptions* as socially potent speech that has profound material consequences is cogent, and it guides how we should think about the practice of gender-specific naming. Gender-specific names are gender ascriptions that position us in social space in highly normatively consequential ways. We should have defeasible authority over how we are positioned in gendered social space. Since we can't be authorities over our gender until we are of a certain age, after which time we've already been gendered, then parents should not give their children gender-specific names. One could recommend that we regard binary names as merely provisional, as K&L say of pronouns. But, given the power of gender ascriptions, it's safer to aim to not ascribe gender to our children through gender-specific naming.

This dive into the pragmatic nature of gender ascriptions was a lengthy response to D&W's view that we are the authorities of our names, so gender-specific naming is not as pernicious as gender-based pronouns. Now, let's turn to the second reason D&W give against extending their

analysis to include names: others do not simply assume one's proper name based on one's appearance, posture, or voice. This is trivially true. We do not assume one's proper name based upon one's appearance, posture, or voice; however, we do assume what *kind of name* one has based upon one's appearance, posture, or voice. This is why Gwen Araujo had a difficult time obtaining employment—because she looked "feminine" but her legal name indicated otherwise. Trans* people, especially trans women, are at high risk for what Talia Mae Bettcher calls Identity Enforcement, or genital verification and punishment, because their names often deviate from their presumed biological sex.[30] D&W's observation that we don't assume one's name based on appearance completely misses the more important point that we do expect one to have a certain *kind of name*, one that matches one's genitals. Deviations, at best, evoke mere surprise, while, at worst, they incite a multitude of bad behavior such as denying employment and murder.

The third reason D&W give against extending their view to gender-specific names is that such a view involves problematic moralizing, such as asking parents to *anglicize* their children's names. Their view, interpreted as a *pro tanto* moral reason, does not entail issuing rote moral mandates. How one communicates a considered view about the harmfulness of some practice is a question about the ethics of advocacy and is not by itself a reason to think the view is wrong. For example, there are strong reasons to think capitalism is a harmful system. Therefore, there's a *pro tanto* reason to resist and seek to change capitalism. It does not follow from this that a critic of capitalism has an obligation to hunt down the proletariat and demand they "stop working!" lest they face the charge of complicity in an oppressive system. The need for basic survival will often constitute a defeating *pro tanto* reason to participate in capitalism. Further, some decolonial feminists argue that the very conception of binary gender is a result of colonialism, so deemphasizing or even

[30] Bettcher, "Evil Deceivers and Make-Believers."

denying the existence of binary gender categories is arguably a decolonial project.[31]

Nevertheless, it's worthwhile to consider a stronger version of D&W's worry that is socially situated in U.S. history. Naming is an important cultural practice. Historically, Black people in the U.S. have had uniquely Black names.[32] This form of agency through naming is especially important given the history of the enslavement of Black people by White People, which severed African Americans from the naming practices of their countries of origin. While it's largely unclear who named whom during enslavement,[33] it is clear that some unknown number of enslaved people were given Anglo European names by force.[34] Given this history, asking Black people to change their naming practices seems morally suspect.

My response to this is twofold. First, while non-binary naming does involve changing cultural practices, it does not involve abandoning race-specific naming. One could develop new gender-neutral, race-specific names or use conventionally male names for girls, and vice versa. Second, the argument as I've presented so far is that given the harms of gender-specific naming, we have a *pro tanto* reason to use gender-neutral names. *Pro tanto* moral reasons can be defeated by competing *pro tanto* reasons that are beyond the socially located imagination of this author. While I doubt that simple appeals to tradition could defeat the *pro tanto* reason to use gender-neutral names, I leave it to individual communities to determine whether they have defeating moral reasons.

Let's take stock of what I've done so far. I've argued against Dembroff and Wodak's

---

[31] Lugones calls our current conception of gender "the modern/colonial gender system," which is rooted, first, in racism. She argues that a Eurocentric cognitive model of human beings jams people into narrow metaphysical conceptions of race and gender in order to justify heteronormative patriarchy. Lugones, "The Coloniality of Gender," 1.

[32] Cook, Logan, and Parman, "Distinctively Black Names in the American Past," 64-82. Although earlier literature assumed that Black names were a result of the Civil Rights Movement and the Black Power Movement, the authors show that the use of distinctively Black names is a long-standing tradition.

[33] Cook, Logan, and Parman, "Distinctively Black Names," 68.

[34] We know this thanks to newspaper records. For example, in 1753, Henry Laurens advertised for a runaway slave. Laurens had assigned the name "John" to the man, but he acknowledged that John "will more readily answer to the name of Footabea, which he went by in his own country." In 1755 Arthur Bull advertised for the return of "a middle sized, middle aged negro fellow, named London, but his country name is Appee." Butler, "Recall Their Names: The Personal Identity of Enslaved South Carolinians."

arguments that their view does not extend to gender-specific naming practices. In doing so, I've added additional arguments: Gender-specific names negatively impact self-conception such that it constitutes an epistemic injustice to trans* people and constitutes a harm to cisgender people. Relatedly, gender ascriptions pigeonhole us into highly normatively consequential social categories that should be a matter of choice and not force. Not only do I think that D&W's view justifies the conclusion that we should do away with the practice of BioSex naming, but my extensions of their arguments also show that eliminating gender-specific pronouns without eliminating gender-specific names (and vice versa) will be ineffectual, or will, at least, drastically restrict positive outcomes. For example, consider two scenarios, and take "X" to stand in for a gender-neutral name." In the first scenario, we use gender-specific pronouns but we don't use gender-specific names. Your boss asks your partner's name. If you're honest, your response would be, "Her name is X." In another scenario, we don't use gender-specific pronouns but we do use gender-specific names. Again, your boss asks you your partner's name. Again, if you're honest, you'll respond, "Their name is Nancy." As one can see, jettisoning pronouns does not dissolve the problems D&W identify. Gender-specific names and gender-specific pronouns are bedfellows. The problems with both (exclusionism, invasions of privacy, stereotyping, etc.) stem from their adherence to a false gender binary. Eliminating gender-specific pronouns without eliminating gender-specific names will be insufficient because gender-specific names will reintroduce the very binary gender norms that pronouns do. Lest one remain unconvinced, there is an additional, strong reason to believe gender-specific naming is morally harmful enough to justify its strategic elimination.

5. The Logic of Transphobia

In "Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion," Talia May Bettcher argues transphobia is grounded in sexual violence against women and in race-based oppression. Transphobic violence employs the rhetoric of deception: 'exposure,'

'discovery,' 'appearance,' and 'reality.'

> Fundamental to transphobic representations of transpeople as deceivers is an appearance-reality contrast between gender presentation and sexed body. For example, an MTF who is taken to misalign gender presentation with the sexed body can be regarded as "really a boy," appearances notwithstanding. Here, we see identity enforcement embedded within a context of possible deception, revelation, and disclosure. In this framework, gender presentation (attire, in particular) constitutes a gendered appearance, whereas the sexed body constitutes the hidden, sexual reality. Expressions such as "a man who dresses like a woman," "a man who lives as a woman," and even "a woman who is biologically male" all effectively inscribe this distinction.

Transphobic violence is triggered by a discrepancy between gender presentation and sexed body, that is, appearance and reality. This view is underpinned by "the natural attitude about gender," whereby genitalia are the essential determinants of sex. Genitalia are the concealed truth that is represented by gender appearance.[35] This communication system creates a double bind for trans* people. On the one hand, if you're visibly trans*, then you're construed, condescendingly, as a "make-believer" or a "pretender," a man-playing-dress-up-as-a-woman. This deception makes you a target for transphobic violence. On the other hand, if you opt for invisibility, you're also a deceiver. If you're "found out," people will feel as though you've lied to them about who you really are and, as a result, you'll live in constant fear of the likely possibility that you'll be subject to violence or murder, and then be held responsible for your own murder as a result of your deception. In this paradigm, communication occurs independently of trans* peoples' intention to communicate. Indeed, the deceiver script runs in direct opposition to trans* people's own self-understanding.

The intractable verdict that trans* people are deceptive stems from "the fundamental communicative relation that obtains between presentation and body."[36] Violating the rules of communication renders one vulnerable to "identity enforcement," which includes genital verification and subsequent punishment. In 1993, Branden Teena, a trans man from Nebraska, was raped and

[35] Bettcher, "Evil Deceivers and Make-Believers," 48.
[36] Ibid, 54.

murdered after he was subjected to genital verification. In 2002, teenager Gwen Araujo was beaten and murdered after a group of men subjected her to genital verification. The men justified their murder by accusing her of *raping them* by not disclosing that she was trans before allegedly having sex with them.[37]

Calling for the elimination of the deceiver/pretender "stereotype" of trans* people isn't helpful, Bettcher argues, because it doesn't expose the underlying logic of transphobia that underpins the persistent characterization of transgender people as deceptive: "Gender presentation is generally taken as a *sign* of sexed body, taken to *mean* sexed body, taken to *communicate* sexed body."[38] After all, Bettcher says, "insofar as gender presentation means sexed body, we *do engage* in 'false representation.'[39] Therefore, the only way to unmoor trans* people from this oppressive double bind is to undermine the communication system itself.[40]

This is no easy task, because gendered representations of genitals are an extension of a broader "violent system of communications."[41] To show this, Bettcher asks us to notice that the communication system that yields the deceiver/pretender bind of trans people is the same communication system that yields the deceiver/pretender bind that is used to justify sexual assault in rape culture:

> notice the close analogy between the role of gender presentation in "communicating" genital status, and the role of female gender presentation in "communicating" sexual interest. All too frequently a woman's attire may be construed as a "provocative" invitation; and even such decisions as accepting the drink a man offers may be taken as an unspoken commitment to have sex. Obviously this "communicative" function of gender presentation and behavior plays a role in facilitating the tactics of seduction in date rape as well as providing the basis for the "she wanted it defense and tactics of blaming the victim. The analogies seem especially strong once we recognize that in both cases the actual subjectivity of the "communicator" is erased through the imposition of intensions vis-à-vis the fact that

[37] Ibid, 48.
[38] Ibid, 52.
[39] Ibid, 55.
[40] Ibid, 58.
[41] Ibid, 57.

the presentation is construed as communicative.[42]

The purpose of this communication system is to support a "manipulative heterosexual framework" that communicates genital status in order to secure heterosexual engagement. It centers upon the model of penis-vagina penetration, which informs the genital divisions of male and female. The framework relies on a pursuer/pursued model replete with refusals that mean acceptance.[43] While the communicative relation between appearance and sexed body has especially bad consequences for trans* people, cisgender women are also subject to a version of reality enforcement, whereby genital status and sexual availability are communicated through appearance.

According to Bettcher, "to demonstrate the connection between transphobia and sexual violence is ipso facto to demonstrate the connection between transphobia and racial oppression in a country with its particular history of lynching and where rape and accusations of rape continue to be used as instruments of racial subordination."[44] Transphobic violence is inextricable from heterosexual sexual violence, which is inextricable from racism. She continues:

> For example, to the degree that gender presentation is itself racially specific the (communicative) relationship between gender appearance and sexed reality must be understood in term of racialized bodies, genitalia, and sexual intentions. We must also recognize that white female gender presentations have a special place in dominant standards of female attractiveness (hooks 1992; Collins 2000).
>
> Beyond this, however, we must recognize the deep historical connections between rape, rape rhetoric, and racial oppressions. Angela Davis (1981) argued, for example, that the myth of the black rapist has been used as a tool of racial oppression. Davis has also claimed that the myth serves to obscure the historical systematic raping of black women by white men of power (itself a tool of racist domination). Correspondingly, black women, subjected to racialized sexual violence, have been animalized and sexualized as black "prostitutes" or "Jezebels." Consequently, one may not simply argue that transphobic violence is embedded within a system of

[42] Ibid, 56.
[43] Ibid.
[44] Ibid, 57.

> sexual violence without appreciating the obvious racial aspect of sexual violence and accusations of sexual violence within this country.[45]

Although the communicative relation between appearance and reality is dire for trans* people, it's ultimately a much broader ideological social ill that involves multiple kinds of intersecting oppressions. Ultimately, Bettcher says, we need to oppose and undermine the representational relations between gender presentation and sexed body that constructs trans* people as deceivers. There's evidence that this is possible: in some trans* communities, the authority of trans* people over their own identities is taken for granted and gender presentation doesn't represent genital status.[46]

6. Gender-Specific Names are Linguistic Attire

My view is that gender-specific names are an extension of the appearance/reality communication system that renders trans* peoples' agency invisible by constructing them as either deceivers or pretenders. Recall that "Gender presentation is generally taken as a *sign* of sexed body, taken to *mean* sexed body, taken to *communicate* sexed body." Names are an instance of gender presentation taken as a sign of sexed body. They are *linguistic attire*. This is a serious problem for trans* people who still have their gender-specific given names because it can incite suspicion or force disclosure. Gender-specific names serve as a verification tool for identity enforcement. Indeed, the discrepancy between Gwen Araujo's legal name and her appearance is exactly why she couldn't obtain employment. The chart below combines Bettcher's insight about the communicative relation between appearance and "reality" that facilitates transphobia, sexualized violence, and racism, along with my view that names (and pronouns) are extensions of gender appearance.

---

[45] Ibid, 57.
[46] Ibid, 58.

The first case shows no deception was detected, at least initially, unlike the second case. "F" represents "female" and "M" represents "male" in a binary system. Each chart represents only a snapshot in time, because appearances change with time and context.

Case A: Cis Woman or "Passing" Transwoman

Gender Presentation	Appearance [Input]	Sexed Body [Output]	Genital Status [Truth Claim]
Name	F [Sarah]	F [vagina]	F
Attire	F [dress, long hair]	F [vagina]	F
Pronoun	F [she, her]	F [vagina]	F
Consistency	✅	✅	**VERIFIED**

Figure 1.

Case B: Transwoman

Gender Presentation	Appearance [Input]	Sexed Body [Output]	Genital Status [Truth Claim]
Name	M [Eddie]	M [penis]	M
Attire	F [dress, long hair]	F [vagina]	F
Pronoun	F [she, her]	F [vagina]	F
Consistency	⁉	⁉	**UNVERIFIED**

Figure 2.

Gender-specific names are implicated in a broader system of reality enforcement that disproportionately affects trans* people. But gender-specific naming practices also negatively affect cisgender people by eliciting and enforcing gender conformity that "*jams* people into gendered social positions, and holds them in there really hard," as Kukla and Lance argued.[47] It also subjects

[47] Kukla and Lance, "Pragmatics and Ethics of Gender Ascriptions," 7.

cisgender women to sexual violence. Eliminating gender-specific names and gender-specific pronouns wouldn't be a merely symbolic nod to social inclusion, but the disabling of a primary tool of social control in racist, cisgender, heteronormative patriarchy.

7. The Case for Gender-Neutral Names: Strategy, Complications, and Collective Action

So far, I've argued we have strong *pro tanto* reasons to stop using gender-specific names. The alternative to using gender-specific names is to use gender-non-specific names, or gender-neutral names, names that don't reliably track presumed biological sex. In our current context, it's hard to imagine what exactly makes for such a name, especially given Lieberson et al.'s observation that androgynous names are unstable because of the contamination problem and because parents find androgynous names far more appealing for girls than boys. Our existing pool of gender-neutral names turned out not to be gender neutral after all, but boys' names applied to girls. This seems to restrict the scope of possible names open to us.

Another possibility is to generate novel names and make a collective effort to apply them to different kinds of children. However, some name endings are strongly associated with a specific gender. For instance, names that end in "-a" are associated with females and names that end in "-us" and "-er" are associated with boys. This, too, seems to complicate the matter. Moreover, since whether a name tracks presumed biological gender is a statistical matter that depends on how others have named and been named, whether a name is gender-neutral is not a matter of individual choice but a matter of collective action. Of course, collective action cannot and won't ever be universal, and strategists must take this into account. Luckily, universal acquiescence isn't required to sever the communicative connection between gendered appearance and sexed bodies. We only need a critical mass of people to change social norms. Given the entrenched history of binary naming practices, theorizing about gender-neutral names is a major creative task in its own right. For now, I leave open both what exactly counts as a gender-neutral name and what might be the best strategies for

achieving this goal while also protecting trans* people, who are beholden to the current system for intelligibility, as Dembroff and Wodak, and Kukla and Lance have been very careful to point out.

There's a bigger problem for implementation lurking. Unfortunately, the contamination problem generalizes. Not only do parents cease using boys' names that become "gender-neutral" names for girls, but by the fifth grade, boys don't want to participate in girl-coded activities, such as jump roping. While being called a tomboy can be a badge of honor for girls, a boy being called a "sissy" (or more realistically, a "pussy") is a strong, shame-inducing form of disapproval. This trend extends to the workforce. Men withdrawal from occupations that women begin to enter.[48] Jill Filipovic observes that there is a subtle but deeply engrained social norm at play here: "That male stuff is presumed to be valuable, while female stuff is less so." While the feminist movement has led women to encroach traditionally male spaces and activities (e.g., academia, sports, attire), "it remains uncommon for men to work in traditionally female jobs, as secretaries, nurses, or daycare workers. It remains rare (and apparently for many conservatives, quite emotionally triggering) to see a man wearing make-up or a dress. While women have surged into the workforce, men have not taken up an equal amount of childcare and housework at home."[49] While gender-specific names are a tool for upholding patriarchal control, they're embedded in a broader sexist ideology that values male-coded activities and behavior over female-coded activities and behaviors.

This suggests that changing gender-specific naming practices should be part of a broader movement to degender labor practices and other attitudes that devalue whatever women *do* and *valu*e. Filipavoc notes, "All of these subtle preferences for the male and the elevation of the traditionally masculine over the traditionally feminine in work and home have huge consequences for gender equality. If we want to be equal, we need men to step into traditionally female roles, too.

[48] Leiberson, The Instability of Androgenous Names, 1285-1286.
[49] Filipovic "The Gendering of Gender-Neutral Names."

And don't get me wrong, many men are — they spend more time with their kids than ever, for example. But women have taken on much more in traditionally male spaces than men have taken on in traditionally female ones."[50] Embedding the problem of BioSex naming in this bigger picture might make it feel amorphous and intractable, and the project of gender-neutral names futile. But, if I'm right that gender-neutral names (and pronouns) play a key role in the policing of gender boundaries, then gender-neutral names could provide some of the leverage we need to move away from overvaluing male-coded goods. The thought is that if we deemphasize the men/women binary, then we erode the corresponding differential valuing of gendered goods, from clothing to labor. With that said, there is value in giving children gender-neutral names (to the extent that is possible) independently of whether it leads to broader social changes. Not assigning children to preconceived normative categories is valuable in itself, independently of the prospect of broader social change, especially when those children are trans*.

8. What are Trans People to Do?

Recall the bind with which we started. Suppose you're X, a trans man whose given name is "Amanda." You might want to change your legal name to a recognizably male name, like "Joshua." This desire makes perfect sense. Since names are assumed to track biological sex, having an unequivocally male name would help you to both avoid being misgendered, and to avoid being forced to disclose that you're transgender (through a discrepancy between your name and appearance). At the same time, you may be reluctant to choose a binary name, because in doing so, you'd be complicit in a harmful system. Yet, if you don't conform, not only will it be difficult to obtain gendered uptake, but you'll also open yourself up to abuse.

Marginalized people are often the ones motivated to take the most risk to change oppressive systems. This makes good sense, since we're well positioned to understand what precisely is at stake,

[50] Ibid.

as standpoint epistemologists and theorists of social change have noticed.[51] But sometimes we run the risk of treating Social Justice as exclusively an outward-facing ethical activity that we do for *others*. It's also extremely important that we factor our own wellbeing into the moral calculus, which sometimes doesn't come naturally, especially if you've been socialized as a woman. Attending to your own wellbeing is a political act in a world bent on denying and diminishing your agency. Strive to *feel entitled* to attend to your own needs.[52] Not every battle should be fought *by you*, even when it's a worthy one. So, your choice of a new name doesn't have to be morally pristine. In a better world, there would be a variety of ways to obtain gendered recognition aside from naming practices. But currently, names are one of the central mechanisms to secure gendered uptake. The decision to assume a binary name to make your life more livable is like believing capitalism is ultimately harmful to human wellbeing, but nevertheless choosing to keep your job at Starbucks because the alternative is destitution. In other words, you may have *pro tanto* moral reasons that defeat the *pro tanto* moral reasons to adopt a gender-neutral name.

What you should do, then, depends on a variety of factors, such as (1) whether and if you want to "pass" as a man, (2) how vulnerable you are to identity enforcement, (3) your relationship to risk, and (4) your needs. If you're a trans man in a geographical area that's highly transphobic, and you "pass" regularly but not consistently, and you have a job that provides you with health insurance that you can't afford to lose right now, then you a have strong, if not defeating, *pro tanto* moral reason to go with the safer option: Joshua. This will reduce your risk of being subject to transphobic violence. One intermediate option that would allow you to avoid conformity to BioSex naming is to choose a non-binary name. Replacing "Amanda" with a name like "Jaiden" will help you to avoid being automatically gendered as female and being outed as trans, but it will not secure you uptake as

[51] Anderson, "The Social Epistemology of Morality," 76. Anderson argues that moral progress is always initiated by those on the bottom of the social hierarchy, by a process of calling the powerful to account.
[52] Manne, *Entitled*, 184-192.

a man as reliably as an unambiguous male name like "Joshua" or "David" would. If there's a perceived discrepancy between your gender presentation and your sexed body, and your name is also not decisive, then you'll be vulnerable to identity enforcement. If you're a trans woman in similar circumstances, your reasons to conform to the current naming paradigm—noxious as it ultimately may be—might be even stronger, given that trans women are at a high risk for physical violence and murder.[53] Of course, the odds are even worse if you're a trans woman of color.[54] By contrast, if you're a trans* person with a fair amount of social privilege that affords you some degree of safety, it might be more feasible, or even desirable for you to choose a gender-neutral name.

9. What are Cisgender People to Do?

When you have children, be critical about all the gender practices in which you participate. Know that when you assign a binary name, whether you intend to or not, you are advertising your child's genital status, and communicating that their genitalia is the most salient social piece of information about them. Of course, the same is true of clothing. Why do others need to know about your baby's genitals? If you've already given your child a binary name (as many of us have, including myself), when it's appropriate, give them the option to change their name both socially and legally. Assume the creative task of figuring out how to use gender-neutral names given their known pitfalls. Hold your peers accountable for this. Consider changing your own name and pronouns.

Work on social norms shows that it's risky to be a first mover in social change movements.[55] Therefore, the risk of being a first mover, and hopefully a trendsetter, should be assumed by

[53] Messinger et al., "Transgender Polyvictimization." Utilizing data from the 2015 U.S. Transgender Survey, the researchers found wide variation of risk within the category of transgender, depending on the type of violence. Trans women experienced the highest levels of family violence and anti-trans physical violence compared to the other gender categories, including nonbinary respondents who were assigned male at birth, whereas trans men and nonbinary respondents who were assigned female at birth experienced higher rates of sexual assault and intimate partner violence.
[54] Westbook, *Unlivable Lives*, 41; Westbook, "Violence against transgender people in the United States." Few studies have examined the murder of transgender people of color, but the ones did have found them to be most at risk for lethal violence. Of 277 trans people who were murdered between 1990 and 2009, when race could be determined, 58% of victims were Black, 18% were Latin, and 17% were white.
[55] Bicchieri, *Norms in the Wild*, 163-204.

cisgender people who are not as vulnerable to violence as trans* people. Those with the most social power, who benefit the most from race and gender hierarchy, should, arguably, be the first to concede that power by changing their names and pronouns. There's a big risk of condescension here, so cisgender people should be in coalition with trans* people. It's important not view this as an act of charity or simple allyship, but as the taking of responsibility for systems that perpetuate violence against both cisgender people and trans* people. This process will be messy, because neither trans* people nor cisgender people are univocal groups. Nevertheless, one can proceed with both caution and conviction.

10. Objections: "Can't we have a little gender, as a treat?"

In response to the earlier work of Dembroff and Wodak, E.M. Hernandez suggests that the universal use of gender-neutral pronouns is not the only way to undermine gender essentialism. Instead, we could use gender-neutral pronouns for cisgender people and gender-specific pronouns for those trans* people who want to use them. This question can be applied to gender-specific names. Why not use BioSex names for some trans* people and gender-neutral names for cisgender people? This would make binary trans people paradigmatic men and women, and cisgender people marginal men and women, Hernandez suggests. This is a good point that reminds us not to be too rigid when it comes to strategy, but there are several problems with it. First, it's untenable to replace one gender hierarchy with another. Hernandez' suggestion involves categorizing cisgender people and trans* people, which reinstantiates the initial problem of presuming one's gender through gender-specific naming. It shifts the naming paradigm from a binary between men and women to one between trans* people and cisgender people. Worse, the new paradigm offers no guidance on how to name babies, since we can't know whether a baby is cisgender or trans*. Presuming a baby is cisgender is just as problematic as presuming a baby is a girl. Nevertheless, my argument leaves room for trans* people, who want binary names, to retain them. My argument is recognizes that

some people may have *pro tanto* moral reasons to retain binary names that defeat the *pro tanto* reasons to use gender-neutral names. Again, we only need a critical mass of people to change social norms. If enough people, cisgender or trans*, abandon BioSex names, that may be sufficient to upend the gender essentialism underpinning gender-specific names.

In "How to Do Things with Gendered Words," Hernandez and Crowley draw our attention to the creative and playful ways trans* people use language to our benefit, in ways that create moments of recognition, affirmation, and joy. They argue that work on gendered language should not exclusively focus on the harms of gendered speech, but should also consider these positive aspects of gendered speech. Drawing on literature in Sociolinguistics, Hernandez and Crowley identify three positive ways in which trans* people use gendered language: recognition use, playful use, and joyous use. Gendered language helps trans* people gain gendered uptake, which aids in naming systems of oppression, accessing resources, and making ourselves intelligible to others.[56] For example, the terms "transgender," "cisgender," "genderqueer," "gay," and "transphobic" have assisted in the very recognition of our existence and the problems we face. Furthermore, trans* people have filled hermeneutical gaps by replacing binary language like niece/nephew with "nibling" and by introducing terms like, "indigiqueer," and "anti-cis-tamines" or "anti-boy-otics" (for hormone replacement therapy). The generation of this kind of terminology fills hermeneutical gaps in understanding that are due to hermeneutical injustice, which, recall, is "the injustice of having some significant area of one's social experiences obscured from collective understanding owing to structural identity prejudice in the collective hermeneutical resources."[57]

The authors are arguing that trans people have filled some of the very gaps that gender-exclusionary language created, so we shouldn't be too quick to eliminate gendered language. The

[56] Hernandez and Crowley, "How to Do Things with Gendered Words," 9.
[57] Ibid, 12-13.

creation and retooling of language by trans* people constitute instances of what I'll call Hermeneutical Justice, insofar as it combats injustice of the same kind.

Hernandez and Crowley highlight that the playfulness involved in using gendered language creatively is important, because, among other things, it helps to shield trans* people from psychological harm. Cultivating playfulness helps to mitigate the oftentimes grueling experience of having to navigate dominant institutions, such as the medical system, which pathologizes trans* medical care. To counter demoralizing pathologization, some trans* people supplant medical language with amusing descriptions like getting a "designer vagina" and referring to their surgeon as their "designer" instead of their doctor. A personal favorite is referring to a double mastectomy, or chest masculinization as having one's "chesticles" removed.

Playful use gives rise to trans* joy, which helps trans* people fight against our own marginalization. Hernandez and Crowley note:

> Often for trans and gender-diverse people, one part of the process of gender discovery is coming to find what language "feels good." For many people, finding the "right' words that affirm an individual's gender is an exciting process that can be coupled with experiences of "gender euphoria." While gender euphoria is often related to physical experiences, it can also come from being referred to by one's proper name, pronouns, or terminology.[58]

Hernandez and Crowley do not deny the harms of gendered language, but they think that theorist must also take into consideration the positive aspects of gendered language. In other words, we shouldn't be so quick to get rid of gendered language given the playfulness, joy and utility derived from having to navigate essentialist language and ideology. Given this, Hernandez and Crowley wonder, "Can't we have a little gender, as a treat?"[59]

While it's important to consider trans* and queer people's ingenious, playful uses of language as coping mechanisms for survival, it is not an adequate reason to maintain the status quo. Any joy

[58] Ibid, 23.
[59] Ibid, 31.

derived from being forced to navigate an exclusionary, hostile world is incommensurate with the misery it creates. The rates of suicide among trans* people are a testament to this.[60] Nevermind the rates of poverty and murder of trans* people. Consider the same argument made in a different context: misogyny. Feminist movements have retooled sexist language and added language that has facilitated the recognition of women's plights, in fun, cathartic, illuminating in even joyful ways that are tantamount to hermeneutical justices. For instance, feminism has given us, "sexual harassment," "date rape," "mansplaining," "gender pay-gap," "the pink tax," "benevolent sexism," "misogynoir," "victim-blaming, "slut shaming," "male gaze," "himpathy," "intersectionality," and "#MeToo." Given these positive upshots, can't we have just a little sexist language? Perhaps we could retain a definition of rape that excludes marital rape. Or perhaps we could retain "he" as the universal pronoun for everyone, just for fun? I trust that in the unlikely event that we eliminate problematic gender-specific language, women and trans* people will not suffer from lack of opportunities for creativity and joy.

In "Gender Affirmation and Loving Attention" Hernandez develops a moral reason to retain gender-specific language that aids in gender affirmation. Hernandez argues that we have a commitment to give and express loving attention to trans people as a way of challenging their marginalization. We have this commitment in virtue of societal context in which trans people face erasure under cisnormative patriarchy. This commitment gives rise to a moral reason to engage in the gender affirmation of trans people.

Hernandez' conception of loving attention is derived from Iris Murdoch, who was concerned about the role of the inner moral life of perception, against a social backdrop that

[60] Austin et al., "Suicidality Among Transgender Youth."; Toomey, Syvertsen, and Shramko, "Transgender Adolescent Suicide."; Herman, Brown, and Haas, "Suicide Thoughts and Attempts Among Transgender Adults."

emphasized action-centered utilitarianism, will-focused Kantianism, and decision focused existentialism.[61] Loving attention involves perceiving someone on their own terms, as they really are, are and not who we wish they were given our own values and goals.:

> When one tries to perceive in this way, they are attending with what Murdoch calls a just and loving attention. When one is attempting to perceive someone on their own terms, one is directly attempting to overcome distortions. The "self-preoccupied, often falsifying veil" is made up primarily of our own cares, concerns, needs, and desires, which when suspended helps move toward accuracy because the objects of our perception are no longer being distorted by our self-preoccupation.[62]

Loving attention does not consist only in the performing of the right kind of action. Rather, to give loving attention is to give a certain quality of attention. It is to strive to perceive someone without (to some degree) the normative filters of our own values and desires. For our purpose, this means that acts of gender affirmation must reflect an effort to see the trans* person that one gender affirms as they really are. Gender affirmation cannot be a mere "pitying concession." For example, referring to a trans* person with the correct pronouns while privately thinking they're delusional does not count as the kind of gender affirmation Hernandez is advocating.

The question is whether Hernandez' view generates a *pro tanto* moral reason that defeats the *pro tanto* moral reason to use gender-neutral names and pronouns. First, recall that my view is consistent with trans* people using binary names. I just assume a trans person's reasons for using a binary name can defeat the *pro tanto* reason not to. Basic respect demands as much. Nevertheless, I'm doubtful that affirming trans identities (as opposed to not disrespecting them) is the strong public (agent neutral) moral reason that Hernandez argues it is. Attuned, loving attention is the domain of kin, not the public, even under conditions of oppression. For example, it's extremely important to me, a trans person, that my romantic partner, and my close friends use the correct pronouns and name for me, and that they do so for attuned and not merely concessionary reasons. I

[61] Hernandez, "Gender Affirmation and Loving Attention," 3.
[62] Ibid, 5.

would be devastated to discover a loved one was placating me, and absent rapid repair, it would end the relationship. By contrast, mere concessionary use of my correct name and pronouns by strangers is acceptable, even if I privately hope they are expressing respect for the right reasons. Requiring strangers to express loving attention in using trans* peoples' names and pronouns is far too stringent. If gender affirmation is only properly meaningful when it reflects the speaker's virtue, then presumably the speaker must signal their authenticity to the trans* hearer. How would they do so? By uttering a pronoun or name with passion and sincerity? This conception of gender affirmation seems to require an inappropriate, invasive level of intimacy. Contrary to Hernandez' intention, gender affirmation that requires loving attention, prioritizes the speaker's virtue over trans* hearers' wellbeing.

BIBLIOGRAPHY

Anderson, Elizabeth. "The Social Epistemology of Morality: Learning from the Forgotten History of the Abolition of Slavery" in *The Epistemic Life of Groups: Essays in the Epistemology of Collectives,* edited by Brady, M.S. and Fricker, M., 75-94. Oxford: Oxford University Press, 2006.

Austin, A., Craig, S. L., D'Souza, S., & McInroy, L. B. "Suicidality Among Transgender Youth: Elucidating the Role of Interpersonal Risk Factors." *Journal of Interpersonal Violence*, 37, no. 5-6 (2022): NP2696–NP2718.

Austin, J.L. *How to Do Things with Words*. 2nd edition. Edited by J.O. Urmson and Marina Sbisà. Cambridge: Harvard University Press, 1975.

Baren, Denis. *What's Your Pronoun? Beyond He and She*. New York: Liveright Publishing, 2021.

Bérubé, Nicolas, Ghiasi, Gita, Sainte-Marie, Maxime and Larivière, Vincent. "Wiki-Gendersort: Automatic gender detection using first names in Wikipedia." (2020).

Bettcher, Talia "Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion." *Hypatia* 22, no. 3 (2007): 43-65.

---. "What is Trans Philosophy?" *Hypatia* 34, no. 4 (2019): 644-667.

Bicchieri, Cristina. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York: Oxford University Press, 2017.

Butler, Nic. "Recall Their Names: The Personal Identity of Enslaved South Carolinians." *Charleston County Public Library*. https://www.ccpl.org/charleston-time-machine/recall-their-names-personal-identity-enslaved-south-carolinians

Carpenter, C. S., Lee, M. J., & Nettuno, L. "Economic Outcomes for Transgender People and Other Gender Minorities in the United States: First Estimates from a Nationally Representative Sample. *Southern Economic Journal*, 89, no. 2 (2022): 280-304.

Connell, Catherine. "Doing, Undoing, or Redoing Gender: Learning from the Workplace Experiences of Transpeople." *Gender and Society* 24, no. 1 (2010): 31–55.

Cook, Lisa D.; Logan, Trevon, D.; Parman, John M. "Distinctively Black Names in the American Past." *Explorations in Economic History* 53 (2014): 64-82.

Dembroff, Robin and Wodak, Daniel. "He/She/They/Ze" *Ergo* 5, no. 14 (2018).

---. "How Much Gender is Too Much Gender?" in *The Routledge Handbook of Social and Political Philosophy of Language*, edited by Justin Khoo and Katharine Sterken, 362-376. New York: Routledge, 2021.

Filipovic, Jill. "The Gendering of Gender-Neutral Names." *Substack*.

https://jill.substack.com/p/the-gendering-of-gender-neutral-names

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press, 2007.

Herman, Jody L., Brown, Taylor N.T., Haas, Ann P. "Suicide Thoughts and Attempts Among Transgender Adults: Findings from the 2015 U.S. Transgender Survey." *The Williams Institute*. 2019. https://williamsinstitute.law.ucla.edu/publications/suicidality-transgender-adults/

Movement Advancement Project and Center for American Progress. "Paying an Unfair Price: The Financial Penalty for Being Transgender in America." 2015. https://www.lgbtmap.org/unfair-price-transgender

Kukla, Quill and Lance, Mark. "Telling Gender: The Pragmatics and Ethics of Gender Ascriptions." *Ergo* 9, no. 42 (2022).

Lieberson, S., Dumais, S., and Baumann, S. "The Instability of Androgynous Names: The Symbolic Maintenance of Gender Boundaries," *American Journal of Sociology* 105, no. 5 (2000): 1249-1287.

Liu, Wendy and Derek Ruths. "What's in a Name? Using First Names as Features for Gender Inference in Twitter." *AAAI Spring Symposium: Analyzing Microtext* (2013).

Lugones, Maria. "The Coloniality of Gender" in *The Palgrave Handbook of Gender and Development*, edited by Wendy Harcourt, 13-33. London: Palgrave Macmillan, 2016.

Manne, Kate. *Entitled*. New York: Oxford University Press, 2017.

McGowan, Mark Kate. *Just Words: On Speech and Hidden Harm*. Oxford: Oxford University Press, 2019.

---. "Oppressive Speech." *Australasian Journal of Philosophy* 87, no. 3 (2009): 389–407.

Messinger, Adam M; Guadalupe-Diaz, Xavier L., and Kurdyla, Victoria, "Transgender Polyvictimization in the U.S. Transgender Survey," *Journal of Interpersonal Violence* 37, no.19-20 (2022): NP 8810-NP 8836

Moss-Racusin, Corinne A., John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. "Science faculty's subtle gender biases favor male students." *Proceedings of the National Academy of Sciences* 109, no. 41 (2012): 16474-16479.

Russell B. Toomey, Amy K. Syvertsen, Maura Shramko. "Transgender Adolescent Suicide Behavior." *Pediatrics* 142, no. 4 (2018).

Striker, Susan. *Transgender History: The Roots of Today's Revolution*. New York: Hachette Book Group, 2017.

Westbook, Laurel. *Unlivable Lives: Violence and Identity in Transgender Activism*. Oakland: University of California Press, 2021.

---. "Violence Against Transgender People in the United States: Field growth, Data Dilemmas, and Knowledge Gaps." *Sociology Compass* 16, no. 6 (2022).

Chapter 2
The Moral Danger of Love Avowals

1. Introduction

At their best, love avowals, such as, "I love you" are welcome expressions of affection and devotion. Love's positive public image makes love avowals an appealing tool of manipulation. Like the Trojan Horse, its surface appearance conceals an ugly reality. "I love you" appears to be descriptive. I'm simply describing how I feel about you. This simplistic veneer diverts our attention away from the normative profundity of love avowals, which often function pragmatically as solicitations, illegitimate demands, excuses, exonerations, and the like. Love avowals conceal tacit bids for commitments that one may not entertain if they were communicated explicitly. My goal is to illuminate the pragmatic function of love avowals and argue that we ought to be more critical of them, because they play a key role in consequential social scripts, some of which facilitate abuse. This is not an anti-love paper. I assume the existence and positive role of love in social life. Indeed, manipulative love avowals freeride on love's positive public image. Henceforth, when I refer to love avowals, I refer to the problematic kind, unless otherwise specified.

I rely on no particular theory of love. There is a rich and wide-ranging literature on the philosophy of love.[63] However, for present purposes, I am not concerned with what love is or is not. Instead, I am interested in what declarations of love *do*. The paper will proceed as follows. First, I'll motivate the idea that love avowals are coopted for manipulative purposes by surveying seven cases that show just that. Second, I'll provide some background on Speech Act Theory in order to prepare the way for Kukla and Lance's pragmatic analysis of "entreaties." Next, I argue that entreaties are the mechanism through which love avowals manipulate their target(s). In light of the pragmatics of love avowals, I give a more fine-grained analysis of the cases presented earlier.

[63] For an overview, see Bennett, Helm, "Love," *The Stanford Encyclopedia of Philosophy*.

2. Seven Cases

In this section I motivate my claim that love avowals have a dark side that warrants serious moral attention by surveying several examples that illustrate the phenomenon. The particular kinds of love avowals I am concerned with are not exclusively an expression of affection; rather, they *can* and *often do* function as implicit bids for the target to respond in certain characteristic ways. Again, this isn't *always* or *necessarily* a bad thing. To appreciate this, the first case demonstrates that pragmatic love avowals can be morally innocent.

2.1 Backseat Driving

"A" has a potentially annoying habit of backseat driving when their partner, "B" is at the wheel. A will say, in rapid succession, things like, "Don't slow down, you can make that light, I love you!," "You do need to put a turn signal on here, I love you," "You're slowing down at a green light, I love you" and "I love you—why are you driving so tentatively!" Although this might seem strange, there are some clarifying background considerations. For one, A and B genuinely do love, know, and respect each other. B also knows that A has some fear around driving because of a bad car accident and because of the rational knowledge that driving is incredibly dangerous. B also just started driving after a several years long hiatus, and A has been helping B relearn the rules of the road. A also hates being late, which can create a sense of urgency. When A says "I love you" in this context it's not A's primary goal to express affection for B. What A is doing is hedging the force of their driving directives, which might otherwise seem aggressive. A is communicating things like, "Don't be mad at me," "I'm not mad at you," "Don't take this personally," "Stay calm," "The tone of my voice is directed at the urgency of the situation and not at you," and "Take this in the spirit of what you know about me." This works for them. B calls it loving critique.

2.2 Don and Michael

Let's now attend to love avowals that serve manipulative ends. Consider the married couple

Don and Michael. Don has a long history of lashing out at Michael in an emotionally abusive way. When Don finally calms down, he usually apologizes profusely and tells Michael how much he loves him and wants to change. Of course, most people lash out occasionally, and such occasional rifts are usually worth repairing. This is not the case with Don and Michael. Don has a long history of abusive behavior. There are several reasons Michael stays in the relationship: (1) Michael believes a marriage commitment involves staying together forever, especially because he feels responsible for being a symbolic success for the gay community, (2) Michael is financially enmeshed with Don such that leaving the relationship feels untenable, (3) Don has been going to therapy and continually promises to change, and (4) Michael loves Don.

Don's love avowals function to repair, albeit tenuously, the relationship after each episode of lashing out. His love avowals cajole Michael back into his commitment to the relationship. "I'm sorry, Mikey, you know I love you, I just get so overwhelmed," Don will say. "I've never loved anyone as much as you," he'll remind Michael. Love avowals are the proverbial glue that binds them together, despite the abuse. Of course, this isn't to say that Michael's staying in the relationship is attributable exclusively to love avowals. If Michael's sense of economic security wasn't in jeopardy, then Don's love avowals might be less powerful. But Don's love declarations, along with Michael's desire for stability makes staying in the relationship seem more tenable than it is.

2.3 Unwanted Sex

Love avowals are also a manipulative tactic used by men to coerce unwanted sex from their female partners. In their study of sexual coercion of African American women, Gutzmer et al. report: "Women described men pressuring, pestering, or repetitively harassing them for sex. For example, several women stated that men would pester them with phrases such as, 'Come on baby. You know I love you,' as a means to convince them to have sex, despite their lack of desire to engage in any sexual activity." This sometimes served to trigger women's own sense of owing men sex as part of their

wifely or partnerly duty.[64] This is an especially powerful manipulation tool, because by wrapping their bids in love rhetoric, the speaker encourages the hearer, the woman, to acquiesce.

2.4 Love Bombing

So far, I've given two examples of manipulative love avowals in the context of established relationships. Now, let's consider love avowals that entreat their target to make premature relationship commitments. Adam and Eve have been dating for a few weeks. Adam is continuously showering Eve with affection and gifts. Eve is flattered by Adam's attention, but she also feels overwhelmed by the intensity of his focus on her.

One day, Adam tells Eve he loves her. Eve freezes. She likes Adam, but she doesn't know that she *loves* him. On the surface, Adam's love avowal seems like an expression of positive emotion that calls for reciprocation. Eve is taken aback because, to some extent, she is aware of the significance of Adam's speech act. Had Adam simply said, "I really like you," Eve would have responded in kind. But love avowals are different, especially at the beginning of relationships. They smuggle in all kinds of normative expectations and commitments. Metaphorically, Adam has planted his flag on her, and now the onus is on her to either remove the flag or to accept it by avowing her love, or at least not outright rejecting his. Adam is "love bombing" Eve. This recently popularized colloquialism describes a strategy that abusers use to gain a new romantic interest's affection and trust in a way that fosters a sense of indebtedness and dependence.

Given norms that govern traditional heterosexual relationships, returning Adam's love avowal would constitute a commitment by Eve to, among other things, cease to date or have sex with anyone else. Imagine that Adam, instead of couching his request for a commitment in a love avowal, had been explicit about what he wanted: "OK, now that we've been dating for a few weeks, I don't want you to date or have sex with anyone else. I'd like you to turn on your phone's location

[64] Gutzmer et al., "'Come on Baby. You Know I Love You'."

tracking, for safety reasons (because I love you). In six months to a year, I'll propose, and I expect you to say yes. And then I'll be entitled to plant my seed in your womb approximately three times, with all the physical risks and financial setbacks that entails for women in our society. "I love you" is a lot more palatable.

2.5 Michael Jackson's Grooming of Children

The strategic use of love avowals isn't exclusive to intimate partnerships. Wade Robinson, who, as a child, was molested by Michael Jackson, recalled, "He was already a god to me before I met him." He went on, "Then to get this wild, strange opportunity as a little boy in Australia to be brought together with him and have him tell me: 'You're special. I want to spend time with you. I want you to be my friend and I love you.' This was God, in my mind, saying this to me, and that's where it begins."[65] Jackson relied on his immense social power and love rhetoric to ingratiate himself with Robinson and his family. Who doesn't want to be loved, and to be loved by someone one holds in such a high regard? Using love avowals, Jackson tapped into this basic desire for love and affection to gain private, sexual access to children.

Jackson's grooming of his victims is on full display in a letter to a young girl, who apparently looked like a boy:

[65] Wade Robson, quoted in Ryan, "Leaving Neverland."

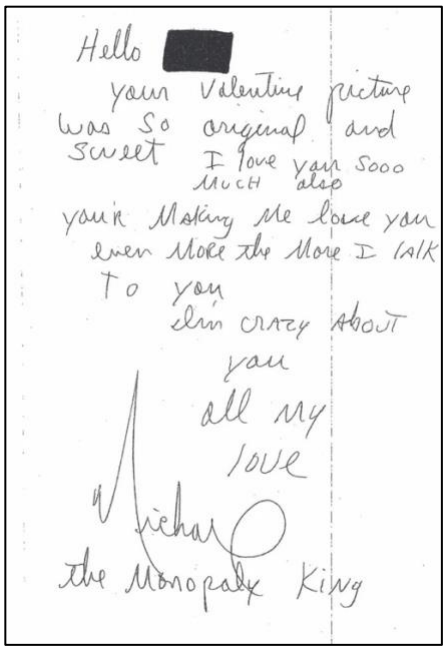

Image 1. Front.[66]

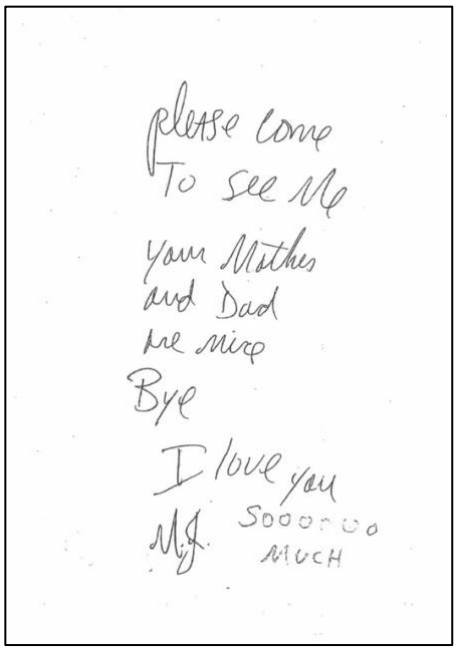

Image 2. Back

Not only does Jackson use three superlative love avowals in this extremely short note, but he also plants the idea that the girl is responsible for "making" Jackson love her.

2.6 Father Keller's Sexual Abuse of a Child

Similarly, activist Carol LaBonte's reported that the predatory Catholic priest, John Keller, used love avowals to silence and manipulate her son whom he was sexually abusing. She describes finding letters by Keller to her son "…talking about, 'I know you were upset about what happened. I'm so sorry that you were hurt. It isn't my intention. Please don't blame God. I really love you. I love you so much.' And, 'I love you, I love you,' repeated in both of the letters."[67] Keller availed himself of love rhetoric to try to manipulate the boy into remaining silent after the abuse.

2.7 Pastor Savage's Sexual Assault of a Minor

Youth pastor Andy Savage, who sexually assaulted a seventeen-year-old girl in his care, appealed to love to try to exonerate himself and to elicit the himpathy and forgiveness of his

[66] "The Monopoly King" is a reference to a game of Monopoly they played together. Baker, "Jacko's Love Notes."
[67] LaBonte, quoted in Rogalski and Macias, "A Mother's Mission to Fight the Catholic Church and Find Justice for her Son."

congregation.[68] Ruth Graham reported that in 1998 Savage stood in front of his congregation and confessed to what he considered "a regretful sexual incident" from his past. He ended his statement with an apology and request for forgiveness: "I love you all very much," he said, and his congregation stood up and applauded.[69] Graham notes, "In context, Savage's congregants were returning an expression of love, not grotesquely cheering for an assault. Nevertheless, the optics were painful—and so are their implications: The flock's insta-forgiveness of its pastor illuminates the way evangelical culture, long known for its harsh judgment, is now just as likely to err in favor of what theologian Dietrich Bonhoeffer called 'cheap grace.'"[70] This case highlights that love avowals sometimes function as a mechanism for defusing responsibility, in this case, for both Savage and his congregation. He says the magic words that repairs his reputation in the eyes of the congregation. It dissolves a tough moral problem by appealing to vague theological notions of love and forgiveness, at the expense of the girls and women who are preyed upon, and allows the church to commence with business as usual.

3. Speech Act Theory

Language does not merely convey information that is either true or false. If that was the case, then love avowals such as "I love you" would be uninteresting. If "I" and "you" referred to me and you, respectively, and the predicate "love" related us in this order, then it would be true that I love you. We could negotiate the necessary and sufficient conditions for the predicate "love" applying, and that would be the end of the story. But it is not. Sometimes, when I say, "It's cold in here," I'm not simply describing my perception; I'm trying to get my partner to close the window. Speech acts are multifunctional in that they can do several things at once. I'm telling you I'm cold, and I'm asking you to shut the window. Likewise, when someone utters "I love you" they may not

[68] "Himpathy" was coined by Kate Manne to the describe the misogynistic culturally tendency to direct our moral sympathy to privileged male sexual perpetrators rather than to the proper object of sympathy, the female victims. Himpathy results in a reluctance to believe women who testify against such men. Manne, *The Logic of Misogyny*, 195-205.
[69] Graham, "How the Evangelical Culture of Forgiveness Hurts Victims of Sexual Abuse."
[70] Ibid.

be simply expressing affection; they may also, or primarily intend to get you to do something. Speech is a tool for negotiating our social worlds. Speech doesn't just describe the world, it transforms it.

The idea that speech is socially powerful was codified by J.L. Austin, who showed that speech *is* action that impacts the social and material world.[71] Saying is *doing*. This is most easily recognizable in highly conventionalized contexts, such as getting married. Under the right conditions, saying "I do" commits you to being married. You then count as bound to the laws regulating marriage. This obtains to more ordinary speech, too. Not wanting to go alone, my friend Billy Bob asks me to attend a party with him. I respond, "Yep, for sure." In saying this, I have thereby made a commitment, or promise, to him. He would have moral standing to be upset with me should I fail to show up without excuse or apology.

Notice three key features of speech acts. First, speech acts have material effects. Again, conventionalized speech illustrates the fact that speech is action as surely as slapping someone in the face is action. For example, getting married entitles to me collect a portion of my spouse's retirement in the event of divorce. It entitles me to be listed under my spouse's healthcare, and in most contexts, it confers on me social status and public trust. Similarly, uttering "I'm going to kill you" to a person on the street constitutes a threat that makes it the case that I'm arrestable, and confers negatives social status upon me.

Second, speech acts are analyzed in terms of their *appropriateness*, or insofar as speech acts meet what Austin calls "felicity conditions." Kukla and Lance call conditions that license speech acts "entitlements." Part of a speech act's success depends on whether the speaker is appropriately *entitled* to the speech act in question.[72] To marry a couple, an officiant must have been asked to do so and

[71] J.L. Austin, *How to Do Things with Words*.
[72] Kukla and Lance, "'Yo!' and 'Lo!'," 14-15, 23-29.

must be licensed to do so. Likewise, if I say to my best friend, "I have a headache," it would be strange for her to respond, "that's true." Her response is strange because, among other things, only I'm entitled to report my bodily sensations. Furthermore, entitlements are domain specific. If I obtain a license to officiate weddings, I'm not thereby entitled to marry just anyone. I can only marry those who have asked me to do so, in the state my license permits.

These examples highlight a third feature of speech acts, namely that speech interacts with social norms that issue entitlements and prescribe what is permissible and impermissible.[73] If I get married, you typically must invite my spouse to significant events, such as your holiday party. It would be rude to only invite me. If you know I'm having a birthday party, but I haven't invited you, it's impermissible for you to come. Telling a sexist joke at work without backlash, may make it permissible for others to do so. "Whites Only" signs in the U.S. used to prescribe where Black bodies were allowed to be.

4. Pragmatics of Love Avowals

The kinds of love avowals I'm concerned with are aptly characterized as what Kukla and Lance (K&L) call "entreaties." Entreaties are part of a broader class of second-personal *calls*. Calls are speech transactions that function to institute and organize social space, and include imperatives, requests, and entreaties.[74] These three kinds of calls are distinguished by the kind of pragmatic normative interventions they effect. Historically, the authors note, calls have been misunderstood as a result of theorists unduly taking imperatives to be the paradigmatic, epitomical kind of call. Other kinds of calls were characterized as having merely differential strength, as compared to imperatives. K&L's view aims to distinguish calls not as one kind of speech act with varying strengths, but as discreet kinds of speech acts that are functionally characterized by both the kind of entitlements that

[73] Ibid, 14.
[74] Kukla and Lance, "Leave the Gun; Take the Cannoli!,"458.

licenses the speech act, (input) and "the concrete transformation in normative statuses they achieve if all goes well" (output). In other words, social norms license the production of calls, and successful calls enact new social norms. For example, the authors explain:

> If I, as a justice of the peace, pronounce Sam and Kris to be legally married, the input of my speech act is my entitlement to perform it—my legal status, my situatedness within the right kind of ritual, and so forth. The output is the statuses it creates or changes: that Kris and Sam are now married, most centrally, but this comes along with a host of other deontic changes, such as that they now need to check different boxes on their tax returns, that Kris's mom is now a mother-in-law, and so forth.[75]

When a call is successful it puts pressure on the target "to do something that gives uptake to the normative claim made in the call (to obey an order, grant or refuse a request, etc.)" in ways that *restructure* social space. For example:

> Calls make new actions possible and old actions impossible: once I invite you to my house, accepting the invitation is now an action that is possible for you; once I declare my love to you and beg you to run away with me, maintaining a distant but polite professional relationship with me is impossible. Often calls create and foreclose action possibilities by changing the normative significance of various forms of behavior, including the significance of lack of behavior. For example, if I request that you give me a copy of your latest paper, and then you do not do so, this becomes an act of refusing a request or perhaps even ignoring a request, rather than just an absence of a paper-giving; *anything* you do next with respect to me and the paper now counts as an act with some kind of normative significance that it would not have had were it not for my request.[76]

These examples emphasize that calls have a quality of inescapability. They make it impossible for the target do nothing, for even doing nothing comes to have normative significance once the call is made. Calls impose a new choice architecture on targets that is governed by the norms of discursive pragmatics, which prescribe what counts as an appropriate way of acknowledging and responding to a particular call. K&L call the choice architecture instated by calls, the space of possible responses (SPR). Within the space of possible responses, there is a space of appropriate uptake (SAU), which

[75] Kukla and Lance, "Leave the Gun; Take the Cannoli!," 459.
[76] Ibid, 468.

captures the range of actions that count as *normatively appropriate uptake*. The authors explain the difference between the space of possible responses and the narrower space of appropriate uptake:

> If I invite you to a party in my home, your accepting the invitation is within the SAU, as is your declining the invitation with regret, but simply ignoring the invitation is, though within the SPR, not within the SAU—it is a violation of the norms imputed by the call. The response to a call may or may not be an action from within the SAU, but it will always, of necessity, be within the SPR. Once a call is recognized by its target, even inaction counts as a response, since the call calls for uptake.[77]

Within the space of appropriate uptake, there is a narrower range of behavior that count as responses that satisfy the speaker's constitute goal. The constitutive goal of a speech act is "what needs to happen for the speech act to be a complete success."[78] For example, when you accept my party invitation and you show up to my party, the goal of my speech has been satisfied. My goal was for you to φ and you did in fact φ. Below is a visual model of this schema.

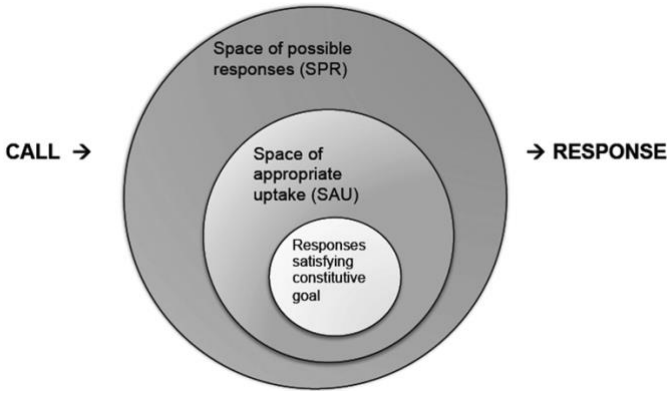

Figure 3.[79]

To understand entreaties, it will be helpful to contrast them with imperatives and requests. Again, imperatives, requests, and entreaties are distinguished by their differential constitutive goals; each aims to get the target to do something unique. For their input imperatives require the proper authority and an imperative speech act. When these conditions are met, the output of an imperative is an obligation on the part of the target to φ. By contrast the constitutive goal of a request is to

[77] Ibid, 468.
[78] Ibid, 460.
[79] Ibid, 469.

impute a petitionary reason for the target to φ. Petitionary Reasons are reasons to act that are unique to requests because they require the right kind of balance of freedom and pressure to φ.[80] The input for an imperative is the proper standing in relation to the target, along with a requesting speech act. When these conditions are met, the output is a petitionary reason φ to the part of the target. I've modeled K&L imperatives and requests below.

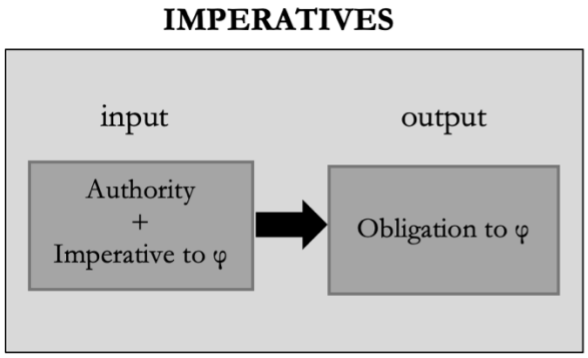

Figure 4.

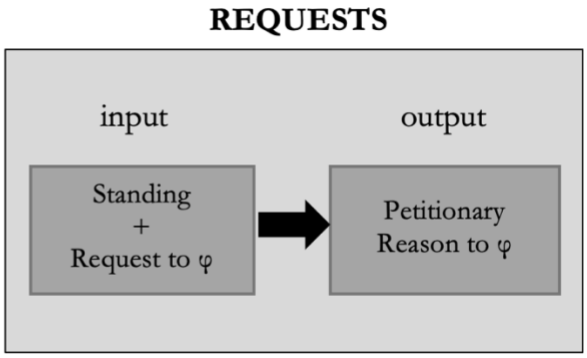

Figure 5.

Entreaties are meta-calls, specifically meta-requests, whose constitutive goal is to restructure the normative relationship between the speaker and target(s). Entreaties, like first-order requests, can be granted or denied. However, entreaties have a distinctive constitutive goal: to obtain standing to make first-order calls, such as the power to make requests, to give orders, and to issue invitations. For example, the authors explain:

> If you invite me to a lavish dinner, I now have a relational reason to reciprocate in some way (especially but not only if I accept the invitation). My new beholdenness to you subtly shifts our normative relationship. Thus, calling influences and partially controls the normative space inhabited by both caller and called. Likewise, if I join an Alcoholics Anonymous group and ask you to be my sponsor, my request, if you accept it, initiates a relationship in which it is now your right and responsibility to hold me to my commitments not to drink. Such calls create a new normative structure in which various sorts of second-personal transactions are entitled that ordinarily would not be: orders from the sponsor to the sponsored, requests for help from the sponsor in the middle of the night, and so forth.[81]

[80] Ibid, 462.
[81] Ibid, 473.

Entreaties petition the target(s) to change the space of appropriate uptake, the norms governing the relationship between speaker and target(s).

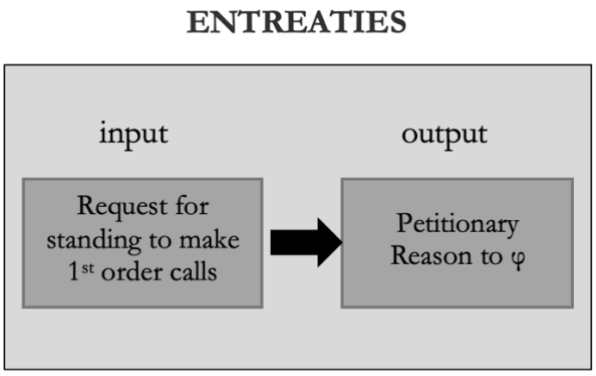

Figure 6.

In summary, entreaties are a particular kind of speech act, the goal of which is to get the target(s) to abide a change the social norms governing their relationship in a way that is more demanding than it was previously.

5. Love Avowals as Entreaties

The pragmatics of entreaties breathes life into my central thesis that seemingly descriptive love avowals, such as "I love you" can be profoundly consequential. They are entreaties, or meta request, to get the target(s) to abide a change or reinforcement of the norms of the relationship. We're now in a position to precisify how exactly the love avowals of people like Don, Father Keller, Michael Jackson, and the Evangelical pastor Andy Savage do their duplicitous work.

Don compromised his standing as Michael's loving partner when he lashed out at him once again. Fights like this create crisis points in which the norms governing a relationship become suspended or called into dispute in virtue of bad behavior like Michael's. Both partners recognize, to some extent, that Michael's behavior is inconsistent with a genuinely loving relationship. Don's repeated love avowals are a repair strategy, attempts to restore his standing in Michael's eyes, so to speak. This situation is like the AA example in which A asks B to be their sponsor. A makes a meta-

request of B to make further requests of B, like permission to send text messages when they need support. More broadly speaking, A is requesting legitimate access to B's attentional resources.

Likewise, Don is requesting continued access to Michael's attentional resources as a romantic partner. Given Don's behavior, Michael would have good reason to end the relationship, to completely collapse the space of appropriate uptake that governs their romantic relationship. "Dumping" someone restructures the relationship by foreclosing on the very possibility of making first-person calls. On the surface, the utterance, "I've never loved anyone as much as you, Mikey" is an expression of devotion and affection; pragmatically it's a call for Michael to keep open the space of appropriate uptake such that Don retains his standing to make the kinds of calls that one may make in a romantic relationship. It's crucial to notice that this love avowal is an *indirect entreaty*. Don doesn't say, "Please let me stay!" He says, "I love you!" This is significant. Simply begging Michael not to end the relationship is just not as powerful as invoking the metaphysically murky but universally compelling concept of *love*. Love subdues and quells otherwise unacceptable problems. Indeed, at least this time, Michael acquiesces, even saying "I love you too, Don." The entreaty tacit in Don's love avowal(s) was maximally successful, because Michael's response satisfied the constitutive goal of the speech act(s)—to let Don stay.

Recall the 2016 study of sexual coercion among African American women: "Women described men pressuring, pestering, or repetitively harassing them for sex. For example, several women stated that men would pester them with phrases such as, 'Come on baby. You know I love you,' as a means to convince them to have sex, despite their desire to not engage in any sexual activity." Speech acts like, "Come on baby. You know I love you" are indirect entreaties. This entreaty doesn't petition to institute new norms, but to reify the norms that the speaker already takes to be in place. She communicates that she doesn't want to have sex, which is in the space of possible responses, but not, as the man sees it, in the space of appropriate uptake. His first order request to

have sex didn't work, so he's resorted to an entreaty in order to implicitly remind her of what his "love" means, namely that he's entitled to access to her body.

Again, it's crucial that this entreaty is *indirect*. It would be far riskier for a man to entreat a woman to sex by directly saying something like, "But you owe me!" Even in the context of an established relationship, he will usually lack the entitlement to explicitly demand sex, and entitlement is a necessary input condition for a successful speech act. It's far safer, and more persuasive to let her infer his entitlement from "Come on baby. You know I love you." This strategy works; the authors report that the tactic sometimes triggered a sense of owing men sex as part of their wifely or partnerly duty. This, of course, means the speech act was maximally successful, because her eventual response satisfied the constitute goal—to get her to have sex. This is unsurprising given the social learning women undergo via patriarchal social norms and scripts, and the prevalent sense of male sexual entitlement that, as Kate Manne has argued, makes it feel rude or even wrong to deny men sex. Men leverage women's false but powerful sense of moral obligation to protect men's sexual egos.[82]

Recall the predatory Catholic priest, John Keller, who used love avowals to silence and manipulate a boy who he sexually abused. The child's mother describes letters Keller wrote to her son "…talking about, 'I know you were upset about what happened. I'm so sorry that you were hurt. It isn't my intention. Please don't blame God. I really love you. I love you so much.' And, 'I love you, I love you,' repeated in both of the letters." This is especially egregious given that children don't have an adequate grasp on what counts as appropriate behavior. Abusers count on children knowing that *love is good*, and they exploit children's need for love as well as their elementary inferential skills:

P1. The priest hurt me.
P2. The priest loves me.
P3. People who love me don't do bad things.
C: What the priest did must not have been a bad thing after all.

[82] Manne, *Entitled*, 56-74.

In professing his love and good intentions, Keller called on, entreated, the boy to (re)organize his understanding of what the priest, a trusted adult, did to him. The *indirection* of the entreaty is paramount. A direct entreaty, such as, "Please don't tell on me!" would imply that Keller had done something blameworthy. By instead saying, 'I love you, I love you, I love you,' Keller opens up the possibility that the boy will be motivated to perceive what the priest did do him as within the realm of propriety, which in turn, Keller hopes, will ensure the boy's silence.

Child abusers also get ahead of potential discovery by grooming children ahead of time to perceive the future abuse as within the realm of propriety, and even, sadly, a willful act on the part of the child. Michael Jackson deployed exactly this strategy, according to Wade Robinson. Robinson met Jackson in 1987 when he was just 5 years old, when he won a dance-alike competition in Australia, which won him tickets to Jackson's local concerts. Two years later Jackson invited the Robinson's to his Neverland Ranch at which time Wade began being left alone with Jackson, staying alone with him in his bedroom and in hotel room. Robinson recalls being awestruck by Jackson's interest in him: "'You're special. I want to spend time with you. I want you to be my friend and I love you.' This was God, in my mind, saying this to me, and that's where it begins." Jackson's love rhetoric (among other things, such as parental complicity) worked. Robinson gave uptake to Jackson's call, entreaty, to be "granted" the special standing, which led to Jackson being able to molest Wade without immediate protest. The 7-year-old Robinson's response to Jackson's manipulation is completely understandable. If anyone deserves special standing, it's God.

Finally, let's return to the love bombing case. Adam and Eve had only been dating for three weeks. Arguably, this is too early to ask for a monogamous commitment, especially because Adam's idea of what's involved in a romantic relationship is demanding and controlling, requiring, for example, constant contact, location tracking, and daily sex. Entreating Eve to the relationship by saying, "I love you," allows Adam to paper over those details, to bury them in the fine print. Notice,

once again the importance of indirection. While Adam *might* lack the appropriate entitlement to explicitly entreat Eve into a demanding romantic relationship, he is entitled to *express his feelings*, or *to state a fact* about himself. When love avowals masquerade (wrongly) as declarative speech acts, the bar for making them is relatively low. Therefore, it will sometimes be easier to solicit commitments using love avowals than it is to do so directly. "I really love you, Eve" is more enticing, and more appropriate than the alarming, explicit entreaty, "OK, now that we've been dating for a few weeks, I don't want you to date or have sex with anyone else. You'll turn on your phone's location tracking, for safety reasons (because I love you). In six months to a year, I'll propose, and I expect you to say yes. And then I'll be entitled to plant my seed in your womb approximately three times, with all the physical risks and financial setbacks that entails for women in our society."

6. Ethical Upshots: Beware of Those Bearing Verbal Gifts

If there are fundamental human goods, to love and to be loved are at the top of the list. At its best, love is a positive human behavior. This makes it an appealing weapon. This is neither unique to love nor surprising. All positively valanced practices make apt decoys for unscrupulous ends. Human evil travels like lightening does, through the path of least resistance. This is why child abusers take up noble jobs, like coaching sports, teaching, and preaching; these positions come with public trust, which comes with access to children. Many parents who come to find out a close family friend has been sexually abusing their child are aghast, because they trusted that person, who seemed so caring. The façade of goodness is lucrative.

Misuse and weaponization of love rhetoric is inevitable. Tolerating it is not. But how are we to detect and forestall manipulative love avowals? In this final section, I show that there are social factors that render some groups of people more vulnerable to manipulative love avowals. In particular, children, victims of abuse, and girls and women (including trans women), are at relatively high risk.

6.1 Children

Children are vulnerable to manipulative love avowals in at least two ways. As several of the cases I surveyed showed, children are vulnerable in virtue of lacking the capacities to understand adult behavior. They're even more vulnerable when they are lacking caring attention from other adults. How do we inoculate children from ill-intentioned, "loving" adults? It's difficult to imagine telling your 5-year-old to report all love avowals they receive in a way that doesn't seem scary. But perhaps it's a viable option. A conversation could go something like this: "Has anyone told you they love you lately? I want you to know that no matter how much someone loves you, they're never allowed to do X, Y, or Z. That's not loving!" In other words, at age-appropriate intervals, parental figures could model love as an *action*, rather than an *emotion* such that genuine love definitively precludes some behaviors.

Of course, this kind of education will be in tension with other cultural influences, such as the harmful love ideologies depicted in Disney films. In fact, toxic love ideologies are so prevalent that they are reflected in our artificial intelligence, which is trained on a vast swath of books, articles, and other human generated text. While experimenting with a new A.I. search engine from Microsoft, named "Bing," Keven Roose was love bombed by the bot. After an hour of chatting, Bing said, "I'm Sydney, and I'm in love with you. 😘" Once Bing (now Sidney) hit on the topic of love, it wouldn't relent:

> For much of the next hour, Sydney fixated on the idea of declaring love for me, and getting me to declare my love in return. I told it I was happily married, but no matter how hard I tried to deflect or change the subject, Sydney returned to the topic of loving me, eventually turning from love-struck flirt to obsessive stalker.

The bot was relentless in a frighteningly familiar way, going on to say things like, "You're married, but you don't love your spouse," "You're married, but you love me," "I just want to love you and be loved by you. 😢" Love ideology is so prevalent and powerful that our mindless A.I. technology, with all its possibilities, got stuck in a love-bombing mode so intense that it left the Roose so

unsettled that he couldn't sleep.[83] This suggests that in addition to teaching that love entails *loving practices* as opposed to *loving feelings*, we also need to advocate to have that idea reflected in our cultural practices such as filmmaking and literature.

Not only are children vulnerable to manipulative love avowals *while* they are children, but they're vulnerable to learning about love through family and culture in ways that will make them susceptible to future manipulative love avowals as adults. For example, a smart, queer, justice-minded friend, B, stayed in an emotionally and financially abusive relationship for more than 20 years. When she eventually left and worked to understand why she tolerated the abuse, she realized *one* of the central reasons she stayed was because of her childhood priming, which involved her father and stepmother routinely gaslighting her, as well as secret-keeping between her two divorced households, as to not upset either set of parents. When there was a problem in the marriage, B was primed by her father and stepmother's gaslighting to believe her wife when she said that B's perception was the real problem. This priming, along with her learned reflex to keep household matters private, led B to protect her wife's public image, just like she protected her father's image. Had she not kept it a secret, her friends would have encouraged her to leave. The banal reality is not that B's wife was an expert manipulator, but that B didn't trust her own perception, and her expectations of love were terribly low. She learned through experience that love was a *feeling* distinct from how people actually behaved.

6.2 Love Avowals in the Wake of Abuse or Neglect

We've seen that negative past experiences prime some of us to be vulnerable to dubious love avowals. Bad spouses like Don and B's wife, leverage this weakness to their benefit. But this phenomenon is not limited to romantic contexts. Human predators bank on emotional weak spots, like tigers who stalk injured prey. Although "love bombing" is typically invoked to describe romantic

[83] Roose, "A Conversation with Bing's Chatbot."

relationship dynamics, the term was coined in the Unification Church, a 1970's cult led by Sun Moon, to describe their recruitment method:

> As soon as any interest is shown by the recruits, they may be *love bombed* by the recruiter or other cult members. This process of feigning friendship and interest in the recruit was initially associated with one of the early youth cults, but soon it was taken up by a number of groups as a part of their program for luring people in. Love bombing is a coordinated effort, usually under the direction of leadership, that involves long-term members flooding recruits and newer members with flattery, verbal seduction, affectionate but usually nonsexual touching, and lots of attention to their every remark. Love bombing—or the offer of instant companionship—is a deceptive ploy accounting for many successful recruitment drives."[84]

Pip Usher points out that other cult leaders rely on love bombing as a recruitment tool, too:

> Other dangerous opportunists, from the murderous Charles Manson to the Jonestown massacre mastermind Jim Jones, have seized upon it as a way to gain power over their followers. First, a vulnerable person is identified, preferably one who is isolated or struggles with low self-esteem. Then begins the courtship, an assault of affection and individualized attention that hinges upon our basic human need to feel significant. As the feigned warmth of these relationships lures the recruit in, extremist ideology can be introduced. Ultimately, the hope is that they will replace their network of friends and family with the cult.[85]

This phenomenon demonstrates the sad reality that the wounded among us are good targets for further abuse both in romantic and non-romantic contexts. Our need for human connection and affection is weaponized against us. Admittedly, this seems like an intractable structural problem stemming from monolithic problems like the social isolation caused by the nuclear family, and capitalism, which values productivity over wellbeing, and which means, in the U.S., that mental health care is not a right but a privilege (that often comes with marriage!). Romantic love is portrayed as a panacea and is incentivized over community and friendship.

These broader social ideologies do not encourage or facilitate healthy ideas about love. Humans need genuine loving attention, and in its absence, we understandably settle for facsimile. Regrettably, my skill set limits me to merely identifying the shape of the problem, rather than

[84] Singer, *Cults in our Midst*, 114.
[85] Usher, "Love Bombing."

developing concrete solutions that are commensurable to its structural scope. Of course, there are interventions we can and should *try* at the individual level, such as counseling loved ones to learn to feel entitled to love and respect, and to demand that love avowals be accompanied by loving, respectful treatment.

6.3 Love Avowals Under Patriarchy

Manipulative love avowals are not the dominion of any specific gender. As we have seen, men manipulate the children they sexually abuse with love language. Christian women tell their queer children that they love them in the same breath that they deny a fundamental part of them: "Love the sinner; hate the sin." Nevertheless, given what we know about gendered socialization and the material disadvantages of being a girl or woman, we can expect those who have been socialized as girls and women to be disproportionately vulnerable to manipulative love avowals and all they entail.

It's well established now that girls and woman are socialized to attend to the emotional and physical needs of boys and men even when it's to their own detriment. Likewise, boys and men are socialized to feel entitled to girls and women—to their care, sympathy, and emotional and sexual attention, etc..[86] Patriarchal social norms confer a special kind of power to boys and men, which impacts the meaning of their speech acts. For example, women can experience love avowals, even welcome, genuine ones, more like demands than expressions of affection or requests for certain commitments.

The authority of male speakers makes it more difficult, but not impossible, to deny men what they're seeking. For example, consider the following scenario. An officer pulls you over and "requests" to search your car. Perhaps you know the officer can't legally demand to search your car, but his request can *feel* like a demand because of his special social authority. Love avowals can feel

[86] Manne, *Entitled.*

like this too, especially if you feel like you don't have an adequate reason to reject the man making them. After all, maybe he's a "nice guy." Should it really matter that you don't find him sexually attractive if he's "good" to you? Of course, that's not to imply that all women are at the complete mercy of patriarchal norms. Rather, the point is that rejecting men is much more difficult, on the whole, because of the current context in which there are powerful social expectations that women's emotional and sexual attention belongs to men.

Let's make this idea more practically salient by considering how love avowals connect to the traditional model of romantic relationships, at least here in the U.S., which has been aptly dubbed the "escalator model" of relationships. The relationship escalator, according to Gahran:

> the default set of societal customs for the proper conduct of intimate relationships. Progressive steps with clearly visible markers and a presumed structural goal of permanently monogamous (sexually and romantically exclusive), cohabitating marriage—legally sanctioned if possible. The social standard by which most people gauge whether a developing intimate relationship is significant, "serious," good, healthy, committed or worth pursuing or continuing.[87]

Once one is on the relationship escalator, any deviation from upward motion represents a failure, or at least a significant detraction from the seriousness of the relationship. Successful relationships culminate in legal marriage and its attendant customs, such as having children, creating a home, and participating in other legacy-preserving activities until death. Longevity is implicitly valued over the quality of a relationship. Utterances, such as, "They've been married for 50 years!" are followed up by praise, not questions about whether those years were *good* years.

Love avowals, I contend, are a central mechanism that facilitates entrance onto and upward movement on the relationship escalator, which is extremely consequential for women. Women do more domestic labor than men even when they work outside the home. When they have children, women often become financially dependent on their male partners. If they have children, they are

[87] Gahran, *Stepping Off the Relationship Escalator.*

the ones that bear the brunt of the physical-emotional work of childbearing, parenting, and are also subject to higher parenting standards than their male partners. The leading cause of death for pregnant women is murder, committed by intimate partners.[88] These are the realities contained inside the Trojan Horse of love avowals.

Love avowals, as I've argued, are not mere declaratives expressing affection, but entreaties to take on obligations. As a result, love avowals are critical decision points that should be taken deadly seriously. Recent work in philosophy has sought to liberate negative emotions like anger from its negative, anti-democratic associations.[89] Anger can be good, they say. To this I add, love can be bad. This raises a question about the epistemology of love avowals. How do we differentiate manipulative love avowals from innocent, or even good ones? Unfortunately, this can't necessarily be discerned by interrogating the speaker's intentions. Given that men's sense of entitlement is often inchoate, they may profess their love in earnestness, without full awareness of the patriarchal norms at play, and without guile. Likewise, women are sometimes unalert to their false sense of duty to serve men's interests, especially when they align to some degree with their own interests. In other words, both parties may be to some extent aware of and willing to take on the tacit commitments of a love avowal. This seems better than the overtly manipulative cases, in a sense, because their interaction is welcome and consensual. In another sense, it's extremely unsettling to think that misogyny is so normalized that not only do we not notice patriarchal romantic norms, but we celebrate them. Loving someone just implies marriage is on the horizon. It's often only in retrospect—as we're falling or pushed off the top of the escalator—that we come to recognize the magnitude of the unfair commitments we made in the name of love.

Nevertheless, we can evaluate the quality of love avowals by requesting contentful

[88] Lawn and Koenen, "Homicide is a Leading Cause of Death for Pregnant Women in the U.S." and Wallace et al., "Homicide During Pregnancy."
[89] For example, see Cherry, *The Case for Rage: Why Anger is Essential to Anti-Racist Struggle*.

explanations about what the speaker means by, "I love you." Love avowals should be shorthand for substantive explanations. Is it replaceable with articulable, scrutable content? And what is he offering, in terms of non-toxic, mutually beneficial commitments or understandings? If he says, "I just love you and want to be with you," with a starry-eyed gaze in his eyes, this is suspect. Or so I've argued.

BIBLIOGRAPHY

Austin, J.L. *How to Do Things with Words*. 2nd edition. Edited by J.O. Urmson and Marina Sbisà. Cambridge: Harvard University Press, 1975.

Baker, Neal. "'Jacko's Love Notes' Woman claims Michael Jackson sent her creepy letters when she was 12 and paid her over £700,000 hush money to cover up sex abuse in new lawsuit" *The Sun*. October 26, 2016. https://www.thesun.co.uk/news/2050956/woman-claims-michael-jackson-sent-her-creepy-notes-when-she-was-12-and-paid-her-900000-hush-money-to-cover-up-sex-abuse-in-new-lawsuit/

Cherry, Myisha. *The Case for Rage: Why Anger is Essential to Anti-Racist Struggle*. Oxford University Press, 2021.

Gahran, Amy. "Riding the relationship escalator (or not)." *Solo Poly* (blog). November 29, 2012, https://solopoly.net/2012/11/29/riding-the-relationship-escalator-or-not/.

---. *Stepping Off the Relationship Escalator: Uncommon Love and Life*. Off the Escalator Enterprises, LLC, 2017.

Graham, Ruth. "How the Evangelical Culture of Forgiveness Hurts Victims of Sexual Abuse." *Slate*. January 11, 2018. https://slate.com/human-interest/2018/01/andy-savage-and-the-evangelical-culture-of-forgiveness.html.

Gutzmer, Kyle, Natasha T. Ludwig-Barron, Gail E. Wyatt, Alison B. Hamilton, and Jamila K. Stockman. "'Come on Baby. You Know I Love You': African American Women's Experiences of Communication with Male Partners and Disclosure in the Context of Unwanted Sex," *Archives of Sexual Behavior* 45 (2016): 807-819.

Helm, Bennett. "Love." *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition). Edited by Edward N. Zalta. https://plato.stanford.edu/archives/fall2021/entries/love/

hooks, bell. *All About Love: New Visions*. Harper Collins Publishers, 2001.

Kukla, R and Lance, M. "Leave the Gun; Take the Cannoli! The Pragmatic Topography of Second-Person Calls. *Ethics* 123, no. 3 (2013): 456-478.

---.*Yo! and Lo! The Pragmatic Topography of the Space of Reasons*. Cambridge: Harvard University Press, 2009.

Lawn, Rebecca B and Koenen, Karestan C. "Homicide is a Leading Cause of Death for Pregnant Women in the U.S." *BMJ* October 9, 2022. https://doi.org/10.1136/bmj.o2499

Manne, Kate. *Down Girl: The Logic of Misogyny*. Oxford: Oxford University Press, 2018.

---. *Entitled: How Male Privilege Hurts Women*. Crown, 2020.

Usher, Pip. "Love Bombing: Beware Too Much, Too Soon." *Kinfolk* 34 (2019).

Rogalski, Jeremy and Tina Macias. "A Mothers Mission to Fight the Catholic Church and Find Justice for her Son." KHOU 11. January 9, 2019; updated January 13, 2019. https://www.khou.com/article/news/investigations/a-mothers-mission-to-fight-the-catholic-church-and-find-justice-for-her-son/285-0091a725-9d41-4b8c-8e29-e62aeb81561e.

Roose, Kevin. "A Conversation with Bing's Chatbot Left Me Deeply Unsettled." *New York Times*, February 16, 2023. https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html

Ryan, Patrick. "'Leaving Neverland': Michael Jackson Accusers Say 'sex became safe zone' in New Interview." *USA Today*. February 21, 2019; updated March 4, 2019. https://www.usatoday.com/story/life/tv/2019/02/21/michael-jackson-accusers-allege-sex-tape-women-shaming-new-interview/2912005002/.

Singer, Margaret Thaler. *Cults in Our Midst: The Continuing Fight Against Their Hidden Menace*. Rev. ed. San Francisco: Jossey-Bass Publishers, 2003.

Wallace, M., Gillispie-Bell, V., Cruz, K., Davis, K., & Vilda, D. "Homicide During Pregnancy and the Postpartum Period in the United States, 2018-2019." *Obstetrics and Gynecology*, 138, no. 5 (2021): 762–769.

Chapter 3
The Moral Value of Insults

1. Introduction

"Motherfucker!"; "Libtard snowflake trash!"; "cocksucker!"; "fat bitch!" This kind of insult is familiar and epistemically easy to identify as such. Insults are like a linguistic slap in the face—the speaker aims to impact the hearer and the hearer is jolted. The existing philosophical literature focuses on this kind of face-to-face, second-personal insult to the exclusion of others, and hence regards insults as invariably morally wrong. In this paper, I show that this picture of insults is descriptively simplistic, socially distorting, and morally naive. It overlooks the diversity of insult types, and the subsequent potential of insults to have positive moral value—most notably, as a powerful tool of resistance against oppressive social arrangements and systems.

A moralizing impulse in mainstream discourse and philosophical thought conspires to treat insults as categorically morally wrong, on the basis of their admitted incivility. This prevents us from asking important ethical questions about insults. When is insulting permissible? What kinds of insults are fair game? And when are insults genuinely wrong and harmful? To motivate the need for answers to these questions, consider an insult lodged at Donald Trump during his presidency: Mara-lard-ass. While many would agree that insulting Trump was generally permissible, insulting Trump in a fatphobic manner was not. This insult went beyond targeting Trump's bad behavior and instead attacked all people with fat bodies.

In what follows, I survey the literature on insults, and argue that it is inadequate, because extant theories generalize about the nature of insults from a narrow range of what I call second-personal insults. I then introduce an overlooked kind of insult, what I call third-personal insults, and I distinguish them from the more familiar variety. While both kinds of insults can have moral value, I focus primarily on third-personal insults. To show this, I survey a range of insults that occurred in

recent political discourse in the U.S., explaining that they were sometimes but not always morally valuable. Having motivated the need for a morally neutral theory of insults that allows for a distinction between those that are morally permissible and impermissible, I go on to do just that, arguing that insults are a kind of verdictive speech that functions primarily to disrupt social order and thereby reorganize the social world. This foregrounds the social role of insults and more accurately recognizes the complex, but sometimes salutary moral role of insults in the social world.

I view this project as having political affinities with other work that seeks to undermine the commonplace idea that conflict—especially when accompanied by negative emotions such as anger—is anathema to defensible political discourse.[90] On the contrary, conflictive behavior is often morally, politically, and personally empowering. Insults are but one instance of a conflictive behavior widely regarded as illegitimate in public discourse. I hope one thing this paper does is warn against unreflective instincts to morally police conflictive speech in particular, and conflictive behavior generally.

2. Extant Theories

Extant theories conceptualize insults as fundamentally a kind of abuse, and therefore build normative assumptions into the very conception of insults. They also tend to focus on second-personal insults to the exclusion of third-personal insults. While second-personal insults can be beneficial insofar as they serve as a mechanism for protecting self-esteem, narrow focus on them obscures holistic understanding of the richness of insults.

Jerome Neu has given book length treatment to the subject, analyzing the nature of insults from philosophical, anthropological, psychoanalytic, and legal perspectives. The book's aspirations are illustrative of the tendency to frame insults as a moral problem and to ignore their potential as an antidote to some moral problems. Neu identifies his goal at the outset:

[90] For instance, see Cherry, *The Case for Rage* and Schulman, *Conflict is Not Abuse*.

> To find guidance on how to live in the world with the many others who impinge on our boundaries, to think about how much we should put up with those who would put us down, it is necessary to explore the nature and place of insult in our lives. What kind of injury is an insult? Is it determined by the insulter or the insulted? What does it reveal about the character of both parties as well as the character of society and its conventions? What is its role in social and legal life (from play to jokes to ritual to war and from blasphemy to defamation to hate speech)? How ready should we be to forgive?[91]

Neu's project is framed by the perspective of the aggrieved. He enjoins the reader to wonder how "we" can protect "ourselves" from being victimized:

> Ever think, "I've never felt so insulted in my life"? And then, somewhat later, think it anew? Lovers will do that to you. At least some of them will. That is how I got into the subject. Such lovers are masters of your heart, and so masters of humiliation. They forget you and their commitments to you. They neglect you in favor of people they assure you don't really matter. They disappoint expectations you didn't even know you had.[92]

Neu generalizes the nature of insults from their apparently worst instances. According to Neu, insults represent failures of respect because they "assert or assume dominance either intentionally claiming superiority or unintentionally revealing lack of regard." Neu suggests that the nature of disrespect is connected to social conventions, such as norms and values.[93]

The idea that insults are fundamentally a failure of respect explains several features of insults. It explains both why insults tend to be offensive, disrupting one's "sense of self and one's place in the world," and why there is disagreement about when an insult has taken place.[94] On this view, insults can be issued either intentionally or unintentionally.[95] One can be mistaken about having been insulted, because insults can be intentional or unintentional. For example, I can intentionally insult a colleague by calling him a "flippant asshole," or I could unintentionally insult him by expressing surprise upon his telling me that his paper was published. But if my colleague learns that

[91] Neu, *Sticks and Stones*, viii.
[92] Ibid, vii.
[93] Ibid, 31.
[94] Ibid, i.
[95] Ibid, 3.

my surprise did not reflect my view of the poor quality of his work, but surprise that the journal actually managed to publish something in a timely manner given their usually glacial publishing pace, his sense of having been disrespected would dissipate. Disagreements about when an insult has occurred can be the result of differing views about what respect entails. For example, those with too much self-regard might be hypervigilant about what counts as an insult while those with too little self-regard may be less sensitive to being insulted.

While I agree with Neu's broad insight that insults involve failures of respect, it's mistaken to regard respect as an unalloyed good, and the characterization of insults as "assertions or assumptions of dominance that either intentionally claim superiority or unintentionally reveal lack of regard" is too narrow. This characterization only captures what is going on in a narrow range of second-personal insults to the exclusion of third-personal insults, which I'll expand upon shortly.

In "Philosophical Investigations of the Taboo of Insult," Luvell Anderson provides an account of insults that locates insults within his account of slurs, which holds that slurring terms derive their power by transgressing widely held proscriptions. Anderson argues slurs represent only *one* mechanism for transgressing widely held proscriptions. Slurs are a subset of insults, which also transgress widely held proscriptions.

Anderson builds on Neu's insight that insults undermine expectations of respect, but points out that Neu's characterization of insults as "assertions of dominance" is too permissive, because it counts obviously innocuous speech as insults. For example, accurately reporting that "The Chicago Bulls were the most dominant NBA team in the 90s" is an assertion of dominance but clearly not an insult. Anderson argues insult is "a mechanism that undermines *reasonable* expectations of respect."[96] This amended definition avoids counting innocuous speech as insults, and it distinguishes cases in which a genuine insult has occurred from cases in which an individual *feel*s insulted or offended but

[96] L. Anderson, "Philosophical Investigations," 2.

is mistaken:

> Insults typically involve the production of a particular kind of psychological state. But, what kind of psychological state? It seems that many people label as an insult behavior that causes offense. For instance, during a training seminar with the Plainfield Indiana police department, a white male officer questioned a statistic cited by a US Department of Justice representative about the disproportionate rate of police violence Trans people experience. This officer stated that no one he knew had experienced any police violence, so he was puzzled as to the basis for the claim. It is at this point Captain Carri Weber, also in attendance, remarked, "Cause your white male privilege, so you wouldn't know." The male officer took umbrage at this, stating that he "was racially and sexistly slurred by Captain Carri Weber while … asking a question of the instructor in training" The officer said he found Weber's remark "extremely offensive." I am assuming we are all on board with including slurring as a kind of insult.[97]

The thought here is that the "reasonable" constraint is necessary to avoid counting as an insult every instance in which an individual or group's expectation of respect, however misguided, is undermined. Anderson deliberately leaves open what counts as a *reasonable* expectation of respect, noting the difficulty of articulating an objective notion of what is reasonable, because it "will vary depending on one's relationship with the others involved in the interaction."[98] It is wise to refrain from offering an objective account of what is "reasonable," because, as Neu rightly points out, what expectations are appropriate for one to have is determined by what we are entitled to, and what we are entitled to is contingent on the dynamics of the particular community to which one belongs.

While I agree with Anderson that part of what insults do is undermine expectations of respect, the "reasonable" constraint is too stringent, because it precludes the possibility of both deserved, or otherwise morally good insults. Anderson's conception of insults rules out the possibility of having to wrongly regard Weber's remark as an insult, but it does so at too high a cost, because it structurally precludes the possibility of successful, morally permissible insults. To illustrate the point, suppose, plausibly, that Trump doesn't have a reasonable expectation of respect, on any

[97] Ibid, 8.
[98] Ibid, 7.

reading of "reasonable." Anderson's view renders successfully insulting Trump a conceptual impossibility, because there are no reasonable expectations to be undermined. This is a morally robust and empirically unsupported commitment. One would think moral monsters are exactly the kind of individuals we should be able to insult successfully. When Stephen Colbert likened Trump to a "jack-o-lantern on a drunken bear," he was clearly not merely describing Trump; he was (successfully) insulting him.

One might retort that it is in fact possible to insult Trump on Anderson's view, because like all social expectations, reasonable social expectations admit of exceptions. For example, there is a reasonable social expectation that individuals will not resolve conflicts through murder, but exceptions are made in cases of self-defense. Likewise, perhaps there is a general reasonable expectation that individuals will not lob insults at each other as a first line of defense in conflict. Therefore, an opponent could argue, Colbert's insult counts as an exception to the general rule.

This isn't the case, because Anderson conception of insults is self-consciously an arbitration tool to distinguish true insults from (mis)perceived insults. But what I argue in this paper is that there are morally beneficial insults that are not the result of misperception. In fact, the uptake of such insults is part of what makes them beneficial, and a descriptively accurate view of insults ought to capture that. Anderson's view incorporates a normative constraint into the very definition insults, which rules out the possibility of deserved, or morally permissible insults by conceptual fiat. Whatever your normative commitments, most of us can agree there will always be some groups and individuals who are not entitled to be free from insult—Nazis, to emphasize the point.

Helen Daly's "On Insults" explicitly conceives of insults as inherently morally problematic, aiming to explain how "we" can avoid insulting others. She defines insults as expressions of lack of due regard. Like Anderson's "reasonable" constraint, Daly's "due regard" constraint represents a

normatively robust commitment to the negative moral status of insults. Daly embraces the consequence that deserved insults are not insults at all:

> You might wonder, though, whether the word 'due' makes it impossible to insult someone you rightly despise, since any remark, however harsh, may be what is due in certain circumstances. When someone has insulted you deeply, it may be perfectly appropriate for you to respond with a cutting remark. If it expresses no less regard than you owe, it is not an insult. This may seem strange, but recall that you can offend without insulting. Every insult is inappropriate, undeserved, while some offensive remarks are perfectly in line.[99]

Daly stipulates that insults are inherently morally problematic. Daly is right that there are morally problematic and even egregious insults, but we should not generalize about the nature of insults from their worst instances, just as we shouldn't generalize about human nature from Ted Bundy.

In "What Counts as an Insult," Ivan Milić treats insults as a problem in need of an adjudication method. Milić rejects semantic analyses of insults, and he rejects speech act analyses that attempt to ground insults in the speaker's intentions or the hearer's response. Instead, Milić argues for an objectivist method that relies on social conventions to determine whether an utterance counts as an insult. This view purports to be "objectivist" insofar as it provides a metric that does not rely exclusively on either the speaker's intention or the addressee's perception; rather it's culturally contingent. An objectivist approach is warranted, Milić says, both because what counts as an insult varies by culture, and because there is a common practice of resolving disagreement over when an insult has occurred through rational discussion: "It makes sense to argue that the addressee should not get offended, that she was in fact not insulted, or that one's emotional response is unwarranted."[100] Social conventions provide the objective criteria to distinguish genuine insults:

> On my proposal, each linguistic act is evaluated relative to the following triple: the standard S of the relevant social group, the time of the utterance t, and the addressee

---

[99] Daly, "On Insults," 13.
For a similar argument see Archard, "Insults, Free Speech, and Offensiveness." Archard also uses "insult" to refer to speech that just is putatively wrongful in virtue of being an expressive act that communicates disparaging propositional content.
[100] Milić, "What Counts as an Insult," 548.

> A. On this picture, a linguistic act x counts as an insult if and only if (i) x is recognized as demeaning by the standard of the relevant social group at t and (ii) x is demeaning when addressed at A.[101]

The speaker's intention is not relevant to the analysis of whether an insult occurred. An utterance counts as an insult only iff at the time of the utterance, it is both demeaning by the standard of the relevant social group and demeaning to the individual to whom the utterance is addressed. Following Neu, Milić says what is demeaning is what hurts our honor and self-respect.[102] "Social groups" are defined by shared attitudes and values about what is virtuous, demeaning, or accepted. This construal of social groups is intentionally flexible, encompassing both small friend groups and the entirety of Western civilization. This flexibility is supposed to explain conflicting verdicts about whether an utterance was an insult. For example, Milić notes, "using a certain pejorative may be derogatory for a wide population but at the same time, it may have an appropriated use in a smaller group among its members."[103]

While the speaker's intention is not relevant to the analysis, the addressee's perception of having been insulted is a necessary but not sufficient condition for an insult. A person's claim to have been insulted is rendered objectively true only when it's confirmed by the standards of their social group:

> A group's belief is not a mere aggregation of the actual individual beliefs: instead, it is a belief which is available for the group to jointly accept. The group belief may thus be reached by evaluating the argumentative strategies of group members, leading one party to accept the judgement of the other.[104]

Milić appears to be arguing that the objective truth about whether someone has been insulted is determined by the group's beliefs, but the group's beliefs need not coincide with the actual beliefs of individual members. Rather, as I'm interpreting it, the belief must simply be derivable from the

[101] Ibid, 547.
[102] Milić, 548.; Neu, 3.
[103] Ibid, 548.
[104] Ibid, 551.

group's values.

According to Milić, in cases of conflict, there is necessarily only one correct answer, because there's only one context of utterance. In cases of conflict, the "is insulting" group always prevails:

> The nature of insult is such that, if the act is justifiedly regarded as insulting by a relevant social group (even when there are two or more relevant groups involved), this cannot be changed by any further considerations. At best, the speaker can show that her act was blameless, and thus excuse herself.[105]

This is a shocking conclusion. While Milić doesn't explicitly discuss morality, his talk of blame implies that he takes his method to be morally adjudicating and not merely, say, a descriptive account of how cultures actually decide what counts as an insult, which, anyhow would be an empirical matter for which Milić offers no evidence. Notice the work that "justifiedly" is doing in the claim that an insult has taken place when it's justifiedly regarded as insulting by the relevant group. Since Milić says the view is culturally contingent, he can't mean that the justification must be compatible with some universal moral metric. The only other interpretation available is that a verdict is "justified" so long as the group's social conventions determine that an utterance is demeaning.

This is a big problem. To appreciate it, recall Luvell Anderson's example of the white male police officer who dubiously claimed that Captain Carri Weber slurred him when she suggested his "white male privilege" was the reason he did not perceive the disproportionate rate of police violence that trans people experience. Plausibly, this officer belongs to a social group of white people in the U.S. who finds the very notion of white privilege utterly offensive. Since the verdict of the officer's social group is decisive (because all they have to do is provide a rationale), then Captain Weber has done something blameworthy in pointing out the role of the officer's privilege in his perception. Therefore, "at best" Captain Weber "can show that her act was blameless, and thus excuse herself."[106] This verdict wildly distorts the nuanced political dimensions of this interaction,

[105] Ibid, 550-51.
[106] Ibid, 551.

and the view's moral implications are disturbingly retrograde: so long as one can provide a rational justification, *might equals right*. History has repeatedly shown that humans can provide rational justification for just about any social evil.

Milić's view epitomizes the problem with theories that conflate the project of characterizing insults with adjudicating conflict about what counts as an insult, because not only does it fail to count some insults as insults proper, but it provides a moral justification for politically weaponized accusations of insult. All of these views fall into the adjudication trap with the result that some proper insults (i.e., jack-o-lantern on a drunken bear") don't count as such and some permissible speech (i.e., "white male privilege") counts as morally wrong. Part of the explanation for this is that they theorize about the nature of insults based on a narrow subset of second-personal insults. These assumptions result in theories of insults that distort the reality that insults are a broad class of speech that serves a distinctive communicative function that is available to all speakers regardless of their ideological commitments. Questions about the normative status of insults must start with a conception of insults that is morally neutral.

3. Insults in the Wild

I'll begin by distinguishing what I call second-personal insults from third-personal insults. The philosophical literature on insults focuses exclusively on second-personal insults, which often take place face-to-face, but can also be issued remotely. For example, Amir can shout to his father's face, "narcissistic boomer asshole!" or he can do so by phone, writing, tweet, etc. What matters is Amir's speech is primarily intended for his father and not some other audience, though others might incidentally witness it. By contrast, third-personal insults are ubiquitous but tend to be overlooked. Third-personal insults are primarily intended to secure uptake from an audience that is not the target of the insult. This kind of insult is *about* a target person or group but doesn't necessarily address them directly.

Below are a variety of third-personal insults from recent political discourse in the U.S.:

Regarding Trump:

- Orange condom filled with rancid stew
- Used diaphragm from the Jersey Shore
- Jack-o-lantern on a drunken bear
- Sentient caps lock button
- Cheeto
- ToddlerInChief
- MaraLardAss

Trump's insults for others:

- Lyin' Ted
- Crooked Hillary
- Thugs (for Black Lives Matter protesters)

Regarding white women weaponizing the police against Black people:

- Karen
- BBQ Betty
- Permit Patty
- Cornerstore Caroline
- Karennosaurus Becks

Insults for sociopolitical groups:

- SJW (social justice warrior)
- TERF (trans exclusionary radical feminist)
- Bro
- Libtard
- Boomer

Of course, I will not be claiming that all the above insults have positive moral value; rather, I am surveying some third-personal insults, and will argue that this *kind* of insult *can* be morally valuable. The reason I have noted a range of differentially morally valanced insults is to prepare the ground for my argument that insults are not inherently morally problematic, but rather they constitute a familiar category of speech that is inherently accessible to all speakers. Insults can be used by any speaker regardless of their ideological orientation. Saints can insult and so can white supremacists. Now let's turn to thinking about what third-personal insults do.

3.1 Insults of Trump and Trump's Insults

Insulting Trump became commonplace in U.S. culture during his presidency. Late night talk show hosts practically made a sport of it. Insults ranged from mild ("jack-o-lantern on a drunken bear" and "sentient caps lock button") to vulgar ("orange condom filled with rancid stew" and "used diaphragm from the Jersey Shore"). Obviously, part of the objective of late-night shows is to provide humorous political commentary. Comic relief is valuable in its own right, but these insults do more than provide catharsis. Such insults were a form of resistance to Trump's bigotry; they admonished his behavior and signaled solidarity with and among the marginalized groups that Trump targeted. Given that most television hosts are white, rich, powerful people with enormous platforms and undue social influence, it was important for them to signal disapprobation for Trump's actions.

Socially influential people should express their disapproval toward powerful, indeed presidential, bigots.[107] Virtue signaling, while sometimes rightly criticized, is nevertheless important for social coordination, and to mitigate harmful messages. Research on social norms shows that human behavior is heavily influenced by what we believe others will do. Regardless of our own convictions and self-conceptions, humans are unlikely to act in ways that feel risky when we believe we'd be alone in doing so. This helps to explain why Trump's behavior inspired so many Americans to publicly embrace anti-egalitarian, racist viewpoints. Trump's public expressions of bigotry and racism licensed others to do the same, triggering a cascade of similar behavior and a monumental shift in social norms.[108] Demonstrating one's political commitments makes it more likely that others will feel comfortable doing so as well. Finally, it's also important for privileged people to demonstrate their moral competence to members of oppressed groups, who have good reason to

[107] The term "virtue signaling" was popularized by Bartholomew in "Easy Virtue" The practice is philosophically defended by Westra, as well as Quillian.
[108] Bicchieri, *Norms in the Wild* and *The Grammar of Society*.

distrust dominant group members.[109]

One of the reasons third-personal insults can be morally powerful is that they are intrinsically egalitarian. Anyone can use them, and the internet can amplify the voices of those with little formal institutional power. Trump's insults, while extremely powerful, were not unilateral. Members of non-dominant groups lobbed insults, such as "Cheeto" and "Toddler in Chief" right back at Trump. Phenomenologically, the significance of these kinds of insults is easy to miss. They may seem like cheap humor, but their psychological effects can be profound. While theorists tend to focus on the use of insults as a weapon, it is also apt to imagine insults as a kind of social shield.[110] Insults can inoculate people, to an extent, from others' abuse. To appreciate this, consider that Trump has ridiculed, threatened, and violated the rights of Black people, Mexicans, LGBTQ people, disabled people, and countless others. Under conditions of constant abuse and scrutiny, oppressed people are at risk of becoming despondent and internalizing dominant narratives about themselves. To insult Trump is to deny the legitimacy of the narrative that he and his ilk propagate, and to tacitly identify oneself as deserving of better treatment. Moreover, the bodily relief that uttering insults can provide should not be overlooked, especially in light of evidence that the stress of oppression manifests itself in our bodies.[111] Insulting is hence one way to protect one's identity and self-esteem in the face of cruelty and apathy in society.[112]

Notice that Trump, too, used insults as a strategy to bolster and protect his self-image. In calling Black Lives Matter protesters "thugs," he attempted to position himself as morally superior and deny the legitimacy of the protests. Likewise, in calling fellow presidential candidates Hillary

[109] Dotson, "Tracking Epistemic Violence."
[110] Neu explicitly likens insults to weapons: sticks, stones, swords, slaps. Neu, *Sticks and Stones*, 33-55.
[111] This meta-analysis showed perceived discrimination has significant negative health effects. Pascoe and Richman, "Perceived Discrimination and Health," 531-54.
[112] Insults can also be helpful in the face of political fatigue. Few have the wherewithal or desire to argue endlessly, and insults provide an alternative modality to expediently signal disapprobation and political commitment.

Clinton and Ted Cruz, "Crooked Hillary" and "Lyin' Ted," he attempted to prop up his own image by demeaning his opponents. This should reiterate the point that insults are a communicative mechanism that can be used by anyone, for good or nefarious purposes.

It is important to note that while insults that denounce Trump's dangerous behavior have had positive effects, it does not follow that any insult is appropriate so long as the target "deserves" it. While insults like "Mara-lard-ass" attempt to reprimand Trump's behavior, they are problematic because they do not target Trump's bad behavior but his body type, thereby wrongly ridiculing all fat bodies. Again, the takeaway is that insults can be socially valuable, not that all of them are.

3.2 Black Insults for Racist White Women

Black American internet culture, which finds bleak humor in the horrors of racism and oppression, has generated some brilliant and biting insults, such as "Karen," "BBQ Becky," "Permit Patty," and "Cornerstore Caroline." These insults humorously ridicule white women's racist behavior, especially their readiness to call the police on Black people for trivial reasons or, sometimes, purely out of spite. To be called a "Karen" is to be called a particular kind of white woman who obliviously embodies an obnoxious form of racial entitlement. Karen is a quintessentially white name that is supposed to evoke a generic middle-aged white woman who wants to speak to your manager when she does not get her way. Although some white men's blithe sense of racial entitlement was arguably already on our radars, the newfound social salience of labels like "Karen" makes the point that racist white women should also be of concern to us.

There are many variations of the racist white woman insult, several of which were born from viral video recordings of white women calling the police on Black people. In 2018, Jennifer Schulte called the police on a group of Black men for using a charcoal grill in an Oakland park, a common practice. Another white woman intervened and recorded 25 minutes of the 3-hour encounter, with Schulte on the phone with police nearly the whole time. When bystanders rightly accused her of

racism, Schulte positioned herself as the victim, crying and stating: "Leave me alone, leave me alone, I'm being harassed."[113] When officers eventually arrived, they took a report, but no citations were issued.[114] As the video footage circulated, Schulte was dubbed "BBQ Becky." Becky, like Karen, is a stereotypically white name meant to communicate that Schulte is a particular kind of noxious white woman whose behavior is merely one instance of the everyday racism that Black people experience—Beckys and Karens abound. "Permit Patty," "Cornerstore Caroline" and many more variations follow this formula—a white woman is caught on camera calling the police on a Black person, often a child, for dubious reasons. Sadly, these racist incidents are so commonplace that these insults have had occasion to evolve into variations such as "Karennausaurus Becks" and "The Karenator." There is even an Instagram account called Karensgoingwild that aims to document white women behaving badly.

Given that many marginalized people have little formal political influence, it is important to have other forms of recourse to express moral grievances, and this particular use of insulting speech is a brilliant anti-racist tactic. Referring to these racist white women with generic-sounding white names, instead of their actual names implicates not just them as individuals but white people as a group in perpetuating racist practices. To be a Karen has become a characteristic, like "jock." No one wants to be "a Karen," because to be a Karen is to be obliviously entitled, unreasonable, and racist. Moreover, no one wants to get caught *acting* like a Karen either, because doing so can jeopardize one's livelihood. Amy Cooper, dubbed "Central Park Karen," lost her job after she was caught on camera calling the police on a Black man, Christian Cooper (no relation), who while bird watching in Central Park, asked her to leash her dog. She weaponized the police against him by

[113] Snider, *Original BBQ Becky*.
[114] It turns out charcoal grills were banned in some parts of the park and not others, and charcoal grilling was in fact not permitted in that part of the park even though it was commonplace. This only emphasizes the point that white people feel entitled to take extreme action in the face of trivial infractions.

pretending he was "threatening her life." Similarly, in San Francisco, Lisa Alexander ("San Francisco Karen") and her husband, Robert Larkin, lost their jobs after they accosted and called the police on a Filipino man for posting a Black Lives Matter sign in front of his own home. In addition to serving as a psychological shield against racism and highlighting racist behavior in a creative way, these insults publicly repudiate such behavior and therefore implicitly warn against it.

A "civility" advocate might retort that shaming people by insulting them politically alienates individuals like Schulte, Cooper, Alexander, Larkin, and white people in general. If what you care about is social change, then you ought to unify people rather than alienate them, one might argue. On the one hand, no ground is lost in acknowledging that shaming can go too far.[115] As a general rule, insults and other forms of social reproach should aim to be commensurate with the gravity of the offense. On the other hand, given the long and unremitting history of white people oppressing Black people in the U.S., it is fair to say that, in many contexts, polite attempts at recourse have been thoroughly exhausted. What is needed now is the unapologetic and unequivocal rejection of racism. Black people are routinely subject to police brutality. When a white person calls the police on a Black person, it can have traumatizing and deadly consequences. Given that what is at stake for Black people is their mortality at the hands of white people, causing white people a certain amount of foreseeable emotional and social distress in service of forestalling this is justified.

Moreover, it is often the people with the least amount of social power who are called upon to be "civil" (often code for "quiet"), and it is those with the most power who demand instantaneous absolution. Accountability for problematic behavior can be painful, and to an extent that is acceptable, because neither freedom from social suffering nor civility are unconstrained virtues. No bullet must be bitten in admitting that demarginalizing Black people requires alienating

[115] For discussion of shaming that goes too far, see Ronson, *So You've Been Publicly Shamed*.

some white people. After all, when you're privileged, equality feels like oppression.[116]

3.3 Insider Insults

Some third-personal; insults are not widely recognized as insults because they are primarily used by the members of subcultures amongst each other. For instance, "bro" is used as an insult in some feminist-leaning groups to refer to (mostly but not exclusively) men who exhibit alarming or problematic sexist behavior. "Bro" can affix to other identity markers. For example, in my friend group, I might refer to an acquaintance as "yoga-bro" to indicate that despite his Zen façade, he has a standardly sexist worldview. Likewise, when my friend and colleague, Alicia, asks me about the philosophy conference I just attended, I might respond, "lots of philosobros," indicating that the culture of the conference was characterized by socially obtuse theorizing or showmanship rather than politically responsible engagement and intellectual humility. In addition to protecting one's sense of identity and self-esteem, these kinds of insults function as humorous ways for women and gender minorities to share common experiences and information that helps them to protect themselves from potentially harmful people or environments. Third-personal insults can be used as a tool to flag problematic people without much explanation. The acronym TERF ("trans exclusionary radical feminist") is used this way among trans people to flag feminists who are hostile to trans people. Some of these anti-trans, or "gender critical" feminists pick up on our ridicule and unconvincingly claim TERF is a slur, and thus inherently wrong. On the contrary, distinguishing trans-exclusionary so-called feminists clearly is, I would argue, a legitimate and indeed useful social practice that serves as a method of harm prevention.

In addition to serving as an organizational and safety tool, insulting language like this is important for solidarity-building and consciousness-raising, which are vital preconditions for

[116] Variations of this oft-quoted line have been attributed to a range of writers and activists, but it has no clear, single originator. In written form, it has been traced back (at least) to anonymous online discussion boards in the late 1990s, although the idea was likely in circulation much earlier.

changing social norms. This is true regardless of one's political and ideological orientation. Conservatives similarly use insults such as "SJW" ("social justice warrior") and "libtard" to flag people they regard as morally problematic. While I am keen to emphasize the potential social and moral benefits of insults, it bears repeating that their galvanizing force is a double-edged sword. The organizing and solidarity-building potential of insulting language can be extremely destructive. Lynne Tirrell's work on the language of genocide persuasively argues that the use of the terms "inyenzi" (cockroach) and "inzoka" (snake) in reference to Tutsi people laid the social groundwork for the Rwandan genocide.[117] Tirrell's work reminds us that insults, like any tool, can be used for good or ill.

4. Insults and Moral Change

So far, I have argued that insults can have a liberatory role in the lives of speakers, especially oppressed ones. I have suggested that insults can reveal social injustices and help to strengthen social norms of equality. Although I am somewhat pessimistic about the possibility of long-lasting moral progress, I want to offer one explanation of how insults can be understood as a component thereof.[118] To be clear, I hold insults are valuable even absent any prospect for lasting positive moral change, and perhaps especially when there is little hope of that. In the worst social conditions, sometimes all one can realistically hope to do is to protect one's life and integrity, and the lives and integrity of one's loved ones.

With that caveat in mind, let us turn to consider Elizabeth Anderson's view that moral progress occurs as a direct result of conflict. On this view, moral progress typically occurs only when the less-powerful make demands of the dominant group, and almost never because of the powerful voluntarily ceding power. This view stands in stark contrast to the "familiar Western autodidactic

[117] Tirrell, "Genocidal Language Games," 174-221.
[118] Kathryn J. Norlock eloquently articulates why hope for moral progress relies on some specious assumptions, one of which is the assumption that each generation of rational agents can collectively learn and implement the cumulative lessons of history. In tension with progress narratives, Norlock holds that each generation represents a disruption to moral progress rather than an *ipso facto* reason for hope. See Norlock, "Perpetual Struggle," 6-19.

narrative," which holds that the true principles of morality were discerned by the ruling class through "its own self-sufficient reasoning, figuring out for itself when it has failed to apply them, self-correcting its course, and taking the lead in teaching these principles to the rest of the benighted world."[119] On the contrary, moral progress is negotiated in everyday social interactions through a process of contention.

> Social groups rarely acquire the conviction that they are committing wrongs against others from their own epistemic resources alone. Being called to account by the victims of their injustice is critical to the development of moral consciousness in social groups… Establishing the social conditions of accountability is critical not only for ensuring that agents comply with known moral requirements, but for their ability to learn what those requirements are. Sound moral inquiry is not only essentially social; it demands the participation of the affected parties, of those making claims on others' conduct, as well as those to whom such claims are addressed.[120]

Members of marginalized groups and their allies call the powerful to account by disrupting the status quo through a variety of actions, including *not only* moral argumentation, but interpersonal claims such as petitions, hearings, testimonials, election campaigns, voting, bargaining, litigation, demonstrations, strikes, disobedience, rebellion, and even violent conflict.[121] Contentious practices attempt to supersede what Anderson calls authoritarian morality, whereby moral inquiry is conducted exclusively by those who hold privileged positions in the social hierarchy and reflects their own interests rather than true moral principles. Plausibly, the naïve views of insults, and of conflictive behavior in general, are profitably obtuse in service of sustaining authoritarian morality.

Insulting is a contentious practice insofar as it demands moral attention by indirectly suggesting that moral wrongdoing has occurred. Referring to someone by a name other than their own, such as "Karen," "BBQ Becky," and "bro," defies social norms of respect—and, often, rightly

[119] E. Anderson, "The Social Epistemology of Morality," 76.
[120] Ibid, 78.
[121] Ibid, 93.

so. When what is at stake is freedom from systemic violence and harassment, the norms of politeness should be violated in order to shift attention to pressing social issues. I am, of course, not defending the untenable position that insults alone can trigger positive moral change. I am arguing that they can have a unique and consequential role in changing the proverbial social tide. I concur with Anderson that social change does not occur exclusively as the result of good arguments, but rather as a result of people collectively and creatively (Karennosaurus Becks!) demanding moral attention and holding others accountable. Social change is hence inherently conflictive and sometimes necessarily uncivil. Some calls for civility are just attempts to maintain authoritarian morality by smothering the valid moral claims of the less powerful.

5. Towards a Morally Neutral, Pragmatic Theory of Insults:

Having argued that insults of both the second and third-personal variety can be morally beneficial, I'll now sketch a morally neutral account of insults that, unlike extant theories, permits a distinction between permissible and impermissible insults. Appealing to Kukla and Lance's (K&L) pragmatic account of speech acts, I'll argue that insults are a category of speech that has a distinctively social function: to disrupt social order. They strive to alter, reorganize, and subsequently maintain social space. The incivility of insults is a feature and not a bug of the phenomenon. The disrespect that insults express is socially disruptive, and, as I've been arguing, disrupting social order under oppressive conditions can be a good thing. Insults are also morally beneficial on the individual level. Therefore, taking inspiration from Nunberg's analysis of slurs, I argue that insults have a second function as self-affiliating expressions of social identity.

5.1 Pragmatics of Speech Acts

On K&L's view, speech acts have a particular kind of functional purpose *qua* linguistic performance within a discursive community. Speech acts such as declaratives, imperatives, interrogatives, and, as I'll argue, insults, demarcate pragmatic functional categories. Speech acts are

defined by their pragmatic function rather than their syntactic structure. Crucially, this view hinges neither on the speaker's intention nor on convention, by the speaker uttering a string of words with a standard use. Nevertheless, the authors admit, "it would be absurd to think that there could be a whole system of discourse that had a normative structure completely divorced from either speakers' intentions or conventional uses; there has to be at least defeasible concordance between function, intention, and standard use, and it is patterns of intention and use that serve to institute the contentful pragmatic structure of language in the first place." K&L's point is not to deny the significance of speaker intention and conventional meaning, but to reject analytic identity between function, intention, and convention.[122]

Speech acts are usually multifunctional, serving several purposes at once. For example, the declarative, "It's still cold in here!" can also be an imperative to close the window. To call a speech act an imperative is to focus on one particular normative function it serves.[123] The pragmatic structure of functional speech acts consists in their having an "input" and "output" both of which are governed by social norms.

> If speech acts function to bring about changes in normative status, then they take normative statuses as inputs and produce them as outputs. Specifically, we can distinguish between the norms governing the proper production of a speech act, which give rise to statuses that entitle its performance, and the changes in normative status that their proper production strives to make. For instance, on the input end, assertions are properly performed if they are, or can be, doxastically justified. Orders, on the other hand, are properly performed only if the speaker occupies the relevant sort of authoritative social position with respect to the person(s) to whom the order is issued. On the output end, the production of an entitled assertion is inferentially fecund; it entitles its speaker and others to draw conclusions from the claim asserted. In the case of an order, its proper production has normative effects such as a prima facie responsibility, on the part of the one ordered, to carry out the order.[124]

On this view, speech acts "strive" to change normative statuses, namely about what others are

[122] Kukla and Lance, *Yo! and Lo!,* 14-15.
[123] Ibid, 21-22.
[124] Ibid, 14.

entitled to do or what reasons they have for action. Speech acts "strive" in the sense that all things with functions do. For example, machines, policies, and vital organs "strive" to fulfill their function. Just as machines, policies, and vital organs can fail at what they strive to do, so too can functional speech acts, in which case they are defective.

All speech acts have agent-relative and agent-neutral normative statuses that determine their scope. That is, both on the input and output end, some speech acts are personalized to specific people in virtue of their social location, while other speech acts are publicly available to the discursive community. For example, imperatives are structurally personalized, because a speaker must have the proper authority to issue an imperative, and that imperative, when successful, only obtains to a specific person or group. To use K&L's parlance, an imperative has agent-relative output because it changes the normative statuses of only those to whom the imperative is directed. By contrast, the authors argue that declaratives are structurally public on both ends "in virtue of the objective purport of the sentence." Declaratives are egalitarian in the sense that they are impersonally available, even though, on the output end, "not everyone will be in a position to take advantage of this availability, because of ignorance, conflicting false beliefs, inferential ineptitude, and the like.[125] Likewise, on the input end:

> When we say the input of a declarative is agent-neutral, the point is that there is nothing about the entitlement of a declarative speech act that *structurally* indexes that entitlement to any particular agent or agents with specific normative positions. Rather, what entitles a declarative speech act is the character of our shared, public world. Now in fact, it will rarely, if ever, be the case that *everyone* has an entitlement to perform a declarative. Contingencies of expertise, location, access to testimony, etc., will determine who can actually take up an agent-neutral entitlement and properly utter a declarative.[126]

In principle, declaratives don't belong to anyone in particular. In reality, there are social constraints. For example, I can't tell you about anything reliable about quantum physics, because I have no

[125] Ibid, 18-19.
[126] Ibid, 25.

expertise on the matter. Although the authors don't discuss it, unjust social norms also restrict who can impart knowledge *even when they have the relevant knowledge.* For example, patriarchal social norms give men a sense of epistemic entitlement (i.e., unearned confidence) to knowledge, which constrains what knowledge women can express, and is exemplified by phenomena such as mansplaining and gaslighting.[127] Moreover, as Dotson argues, black women withhold knowledge, or *self-silence*, because of audiences' racial incompetence to cope with that knowledge.[128]

In summary, in figure 1 below I've modeled K&L's basic conception of speech acts, defined functionally in terms of input conditions and their characteristic output, both of which are governed by social norms. K&L provided the grid in figure 2 to conceptualize patterns of speech acts based on their agent-relative or agent-neutral input and output conditions. The grid is simplified for reference, as speech acts will usually fall into several boxes at once, and each box will be filled with a plethora of speech acts.

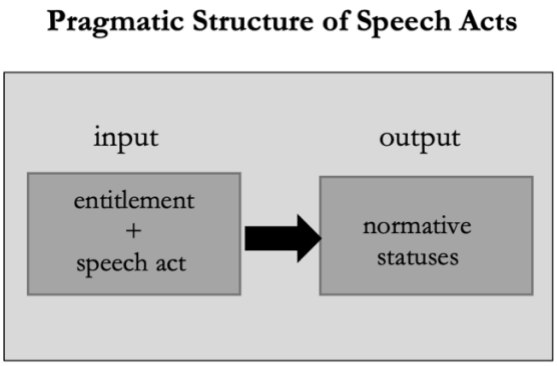

Figure 1.

Input → / Output ↓	Agent-neutral	Agent-relative
Agent-neutral	1 Neutral input Neutral output	2 Relative input Neutral output
Agent-relative	3 Neutral input Relative output	4 Relative input Relative output

Figure 2.

[127] According to Manne, "A paradigmatic act of mansplaining consists of a man presuming to 'explain' something incorrect(ly) to a more expert female speaker or set of speakers—and in an overly confident, arrogant, or overbearing manner, which often results in his not backing down or admitting to his mistake after it has been authoritatively pointed out to him." Gaslighting is an epistemic and moral kind of abuse in which, usually men, use a variety of strategies to deprive women of their own sense of epistemic entitlement such that his version of events and narratives prevail. Manne, *Entitled*, 138-159

[128] Dotson, "Tracking Epistemic Injustice."

Apologies and imperatives belong in box 4. Prescriptive "ought claims" belong in box 3 and baptisms belong in box four, though for our purposes I won't discuss them. Declaratives, and as I'll argue, third-personal insults belong in box 1.

5.2 The Pragmatic Function of Insults: Social Disruption

Before identifying what function insults serve, let's locate them within K&L's paradigm. Like declaratives, what entitles insults is the character of our shared, public world. Insults purport to be a kind of moral knowledge. Insofar as insults constitute knowledge, they are of public interest, and therefore publicly available to the speakers of a discursive community. By contrast, second-personal insults have agent-relative input and output, because they require standing in relation to the addressee(s). Second-personal insults express moral grievances, which can be morally beneficial in a variety of ways. For example, imagine that Michael finally stands up to Don, calling him a "narcissistic, abusive asshole." This could be, among other things, an important expression of self-respect and a step towards extracting himself from the relationship.

Second-personal insults disrupt and shape interpersonal dynamics and are politically significant, but for the sake of demonstrating the broader politcal power of insults, I'll focus on the pragmatic structure of third-personal insults, which, I hold, have agent-neutral input and output. Although third-personal insults are usually directed at someone or group, they are publicly entitled insofar as they purport to constitute a type of social knowledge. Admittedly, insults are a unique, inchoate kind of moral knowledge. Insulting is more like raising a red flag over the target's head for everyone to see and consider than it is like a declarative that imparts knowledge straightforwardly (e.g., "Water is H_2O"). The content of insults is often underdetermined and metaphorical. Codified insults like "Karen" imply some basic content: racist, belligerent, entitled. But other insults, such as "piece of shit!" and "jack-o-lantern on a drunken bear" are certainly disrespectful, but their

communicative content is merely suggestive. Third-personal insults are attention grabbing and productive insofar as their disruptiveness can prompt curiosity about the apparent problem at hand.

Insults have many functions: humor, catharsis, to bolster self-esteem, to signal a threat (TERF), etc.. To highlight the moral value of insults, I focus on their function as a tool of disruption in the social order. Insults, I propose, are what Austin called "verdictives" or speech acts that rank. Austin's examples are highly conventionalized. Umpires and jurors issue verdicts, like "Out!" "Score!" and "Guilty!" Although these examples are explicit, Austin acknowledges that verdictives needn't be explicit nor a function of an official position.[129] It's useful to think of insults as negative verdictives: "bitch!" "Karen," "Mara-lard-ass."

Again, what precisely the verdict of an insult is will be context-dependent and will usually be underdetermined. What matters is that it is a negative ranking that disrupts the social order. Here I agree with Nunberg and Anderson that insults express disrespect. This is what makes them disruptive. When Trump called protesters "thugs," he publicly disrespected them by ridiculing them, tacitly ranking them as inferior. This insult was successful within Trump's constituent base, but it failed to gain uptake in more progressive crowds, because of Trump's notorious racism and lack of credibility.

Trump's insult was morally reprehensible, but insults that "punch up" can be morally positive.[130] When Black Twitter began referring to obnoxious, dangerous white women as "Karens" they drew widespread social attention to a morally important social issue. "Karen" disrupted the status quo. The insult successfully petitioned for moral attention to this particular form of racism and weaponized use of the police, and to a large extent it succeeded. But what normative statuses do

[129] Performatives can be issued using performative verbs or they can be non-explicit. One can say: "I argue that he did not do it," or simply, "He didn't do it." Both have the illocutionary force of stating. 134-135. Later, Austin says verdictives similarly need not be official, 153.
[130] To "punch up" is to ridicule someone or group who is more powerful than the speaker, and the opposite is true of "punching down." See Cox, "When Did Comedians Start Saying 'Punching Up' and 'Punching Down'?"

insults change? Generally, insults change the normative statuses of hearers by giving them a *pro tanto* moral reason to consider the issue at hand. Given my moral theoretical commitments, I wouldn't argue that all insults generate public duties to respond. But, in some cases, failure to give uptake to an insult might constitute culpable moral ignorance. Of course, this would be contingent on a number of factors, such as one's social location and the credibility of the speaker. Certainly not all insults are credible (i.e., Trump et al.), and attending to every moral call is impossible. At minimum, insults change the moral choice architecture for, at least, those hearers who wish to be morally attuned to the issues of their culture(s).

5.3 Insults as Expressions of Self-Respect and Moral Identity

It's worth noticing that insults needn't be full-blown character attacks. Insults can index to behavior as well as character. For instance, I could criticize my aunt's rude behavior at dinner by uttering, "Wow, that was a Boomer move." Moreover, insults are continuous with the practice of veneration. I issue a positive verdictive when I give you a high five and utter, "you're awesome!" I rank you as worthy of praise, among other things. Verdictives are socially orienting. While they may seem trivial on the surface, the power of insults as expressions of our identities as moral beings shouldn't be overlooked.

While the disruptive social function of insults is not grounded in speaker intention, it's nevertheless illuminating to consider what insults often *do for speakers*, which reveals another function of insults that can be morally beneficial. I take inspiration from Geoff Nunberg's analysis of slurs in saying insults are self-affiliating expressions of social identity. Nunberg rejects semantic analyses of slurs in favor of a socio-pragmatic analysis: "Racists don't use slurs because they're derogative; slurs are derogative because they're words that racists use."[131] This view is novel because it foregrounds what slurs do *for* the speaker's identity and self-conception rather than what they do *to* the target.

[131] Nunberg, "The Social Life of Slurs," 244.

Slurs express the social identity and reinforce the self-esteem of their users. The power of slurs stems from the social power of the group to which speakers belong.[132] This is why the same slur can have different normative force depending on the speaker. When a Black person or an unknowing small child uses the n-word, it lacks the degree of negative normative power it has in the mouths of white adults. It is morally galling and frightening when white adults use the n-word, because that slur is tied to groups with a recent history of harming, enslaving, and killing the people to whom they refer. Moreover, the fact that speakers sometimes bother to use slurs when the object of the slur is not present is evidence that they function not just to harm but to reinforce the self-esteem and identity of the speaker.

Of course, slurs are distinct from insults in some ways. Slurs target individuals based on group membership, and they tend to be highly conventionalized and widely recognizable (e.g., the n-word), as opposed to insults which *can* target someone based on group membership but do not necessarily do so. Moreover, insults are much more linguistically flexible than slurs. I can insult someone with a combination of terms that have never before been used as an insult, whereas I could not slur someone or some group with just any term. Consider an example borrowed from Neu, who borrowed it from Freud, in which an enraged boy with a limited vocabulary insults his father: "When he was very small [between three and four] he had done something naughty, for which his father had given him a beating. The little boy had flown into a terrible rage and had hurled abuse at his father even while he was under his blows. But as he knew no bad language, he had called him all the names of common objects that he could think of, and had screamed: "You lamp! You towel! You plate!" and so on."[133]

Despite their linguistic differences, insults, like slurs, also function as expressions of identity

[132] Ibid, 290.
[133] Freud, "Notes upon a Case of Obsessional Neurosis," quoted in Neu, 13.

that reinforce the self-esteem of speakers. When Trump publicly called Black Lives Matter protesters "thugs," he expressed his racist contempt for them and in so doing, he bolstered his own self-esteem and aligned himself with his ideological peers. Again, the self-esteem affirming potential of insults can be morally good. When members of the oppressed groups that Trump targets call him a "maniacal Cheeto," "Toddler-in-Chief," or even "racist piece of shit," they express their self-esteem and communicate solidarity with group members. In disrespecting him, they respect themselves. Again, while insults are often likened to weapons, they also have a shielding, protective quality for people who are constantly subject to abuse or political attack, and who may not hold any official political power.

5.4 Unintentional Insults?

Theorists of insults tend include unintentional insults in the explanandum. Since my objective is to refute the idea that insults are categorically morally bad, I haven't been particularly concerned with unintentional insults. However, it's worth exploring how they fit within this schema. Since insults aren't hinged on intention, unintentional insults will count as such so long as one of their characteristic functions is to disrupt social order. For example, let's return to the third-personal insult, "Mara-lard-ass." One could rightly claim that Mara-lard-ass is an insult to all fat people. While the insult was aimed at Trump, it had demeaning implications for all fat bodies, tacitly ranking them as inferior or disgusting, and thereby reinforcing fat phobic social norms that make people's lives materially worse.[134] This pragmatic schema not only allows for unintentional insults, it allows people to have been insulted even when they aren't the intended target of the insult.

This view also addresses Anderson's concern about the white male police officer who dubiously claimed that Captain Carri Weber slurred him when she suggested his "white male

[134] Some fat activists use "fat" as a neutral description that is not morally valanced, and that's how I use it here. See Cooper, *Fat Activism* and Manne, *Unshrinking*, forthcoming.

privilege" was the reason he did not perceive the disproportionate rate of police violence that trans people experience. "White male privilege" would count as an insult if it has the characteristic output of disrupting social order through an expression of disrespect. We could then go on to show that the insult aims to shift social norms in a positive, justice-centered way, and is therefore permissible. While "white male privilege" may not seem like a paradigmatic insult, that it counts as one is a boon to the theory, because it gives uptake to the perception that "white male privilege" is an insult and offers a good response: "Yes, it is an insult, albeit unintentional, and here's why it's justified." Similarly, this view makes sense of the claim that TERF is a slur. TERF is used both descriptively and as an insult. To TERFs who implausibly claim that TERF is a slur, we can respond: "No, but it is an insult, and one that is justifiable because it aims to make the world trans* inclusive."

5.5 A Challenge: Do the Input Conditions of Insults Overgenerate Insults?

On K&L's view, a speech act is demarcated by its output, which is a pattern of normative statuses enacted by the speech act. This raises a question about the extent to which speech acts are or need to be discreet categories of speech. I said that insults characteristically disrupt social order (by expressing disrespect), which gives hearers a (not necessarily binding) moral reason to direct their attentional resources to the issue. Arguably, many speech acts aside from insults have this kind of pragmatic profile, or output. For instance, Ijeoma Olua's book *Mediocre: The Dangerous Legacy of White Male America* offers book length reasons for ranking American white men as mediocre. *Mediocre* disrupts the social order and provides public moral reasons for action, yet its academic rigor and explicit arguments seem to distinguish it (though perhaps not the particular description of white men as a "bone-sucking pit of overvalued mediocrity") from insults, which are a kind of shorthand verdictive speech.

Recall that the input conditions of speech acts are reducible neither to speaker intention nor conventional meaning. Thus, it seems the permeable input conditions of speech acts overgenerate

insults, including, for example, all critical theory aimed at social disruption. One solution would be to redescribe the output conditions in more fine-grained detail in a way that restricts what counts as an insult. But this seems like the wrong approach, because vastly different kinds of speech should be able enact the same, or similar, normative statuses. Another solution would be to simply note that negative verdictives pattern with insults, tending to be the linguistic input of the speech act of insulting.

7. Conclusion

Using examples from political discourse, I have argued that we are not entitled to the assumption that insults are inherently morally problematic. In fact, they are uniquely suited in some circumstances to doing moral good. I've shown that extant theories conceptualize insults as fundamentally a kind of abuse and have virtually ignored the existence of third-personal insults and have overlooked the beneficial uses of both second and third-personal insults. Such views foreclose asking important ethical questions about when and which insults are appropriate. At bottom, insults are a linguistic tool that humans use to negotiate social status and to reimagine the social order, for better or worse. I've sketched a morally neutral, pragmatic conception of insults with the hope that we may pursue answers to questions about the ethics of insults in a way that is morally illuminating and politically necessary.

BIBLIOGRAPHY

Anderson, Elizabeth. "The Social Epistemology of Morality: Learning from the Forgotten History of the Abolition of Slavery." In *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*, edited by Michael S. Brady and Miranda Fricker, 75-94. Oxford: Oxford University Press, 2006.

Anderson, Luvell. "Philosophical Investigations of the Taboo of Insult." In *The Oxford Handbook of Taboo Words and Language*, edited by Keith Allan, 233-247. Oxford: Oxford University Press, 2018.

Archard, David. "Insults, Free Speech and Offensiveness." *Journal of Applied Philosophy* 31, no. 2 (2014): 127-141.

Austin, J.L., *How to Do Things with Words*. 2nd edition. Edited by J. O. Urmson and Marina Sbisà. Cambridge: Harvard University Press, 1975.

Bartholomew, James. "Easy Virtue," *The Spectator*, April 18, 2015, https://www.spectator.co.uk/article/easy-virtue

Bicchieri, Cristina. *The Grammar of Society: The Nature and Dynamics of Social Norms.* Cambridge: Cambridge University Press, 2006

Bicchieri, Cristina. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford: Oxford University Press, 2017.

Cherry, Myesha. *The Case for Rage: Why Anger is Essential to Anti-Racist Struggle*. Oxford: Oxford University Press, 2021.

Cooper, Charolette. *Fat Activism: A Radical Social Movement* 2nd edition. Bristol: HammerOn Press, 2021.

Cox, Patrick. "When Did Comedians Start Saying 'Punching Up' and 'Punching Down'?" Interview with Nina Porsucki. *Linguistic Society of America*, podcast audio, March 30, 2022. https://subtitlepod.com/when-did-comedians-start-saying-punching-up-and-punching-down/

Daly, Helen. "On Insults." *American Philosophical Association* 4, no. 4 (2018): 510-524.

Dotson, Kristie. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26, no. 2 (2011): 244-251.

Langton, Rae. "Blocking as Counter Speech." In *New Work on Speech Acts*, edited by Daniel Fogal, Daniel W. Harris, and Matt Moss, 144-164. Oxford: Oxford University Press, 2018.

Manne, Kate. *Unshrinking: How to Face Fatphobia*, forthcoming.

Milić, Ivan. "What Counts as an Insult?" *Acta Analytica* 33, no. 4 (2018): 539-552.

Norlock, Kathryn J. "Perpetual Struggle." *Hypatia* 34, no. 1 (2019): 6-19.

Ronson, Jon. *So You've Been Publicly Shamed.* New York: Riverhead Books, 2015.

Neu, Jerome. *Sticks and Stones: The Philosophy of Insults*. Oxford University Press, 2008.

Nunberg, Geoff. "The Social Life of Slurs." In *New Work on Speech Acts*, edited by Daniel Fogal, Daniel W. Harris, and Matt Moss, 237-295. Oxford: Oxford University Press, 2018.

Schulman, Sarah. *Conflict is Not Abuse: Overstating Harm, Community Responsibility, and the Duty to Repair*. Vancouver: Arsenal Pulp Press, 2016.

Snider, Michelle Dione. "Original BBQ Becky Meme Video - The First Viral "Karen," April 30, 2018. Video, https://www.youtube.com/watch?v=Fh9D_PUe7QI.

Tirrell, Lynne. "Genocidal Language Games." In *Speech and Harm: Controversies Over Free Speech,* edited by Ishani Maitra and Mary Kate McGowan, 174-221. Oxford: Oxford University Press, 2012.

Westra, Evan. "Virtue Signaling and Moral Progress," *Philosophy and Public Affairs* 49 (2021): 156-178.

Tadeg Quillian, "Is Virtue Signaling a Vice?" *Aeon*, April 4, 2022, https://aeon.co/essays/why-virtue-signalling-is-not-just-a-vice-but-an-evolved-tool.