The Harms of the Internalized Oppression Worry

Nicole Dular and Madeline Ward

ABSTRACT: In this paper, we locate a general rhetorical strategy employed in theoretical discourse wherein philosophers argue from the mere existence of internalized oppression to some kind of epistemic, moral, political, or cognitive deficiency of oppressed people. We argue that this strategy has harmful consequences for oppressed people, breaking down our analysis in terms of individual and structural harms within both epistemic and moral domains. These harms include attempting to undermine the self-trust of oppressed people, reinforcing unjust epistemic power hierarchies, undermining the moral agency of oppressed people, and obscuring real issues of oppression. In light of these many harms, we conclude that there are pro tanto reasons not to employ this strategy, and instead offer a series of suggestions for those considering adopting such a rhetorical move.

0. Introduction

Standpoint epistemologists typically hold the epistemic advantage thesis, which claims that oppressed people tend to have some epistemic advantages related to knowing about the circumstances of their oppression, because oppressed people have experiences relevant to those circumstances and reasons to think critically about them.¹ And anyone who works on standpoint epistemology has probably heard the following response to the epistemic advantage thesis:

¹ We would like to thank the audience at the 2022 Central Division APA meeting as well as anonymous referees for helpful comments on earlier drafts of this paper. The authors contributed equally to the research and production of this manuscript.
Given the oppressed’s educational deficit, susceptibility to misleading appearances, and motivational pressures towards faulty beliefs, should we think that the [epistemic advantage thesis] is true at all?\(^2\)

It’s a familiar refrain: one introduces a claim about how oppressed people are epistemically advantaged, and one’s interlocutor (or referee, or conference talk audience member) replies, well, how can oppressed people be advantaged since internalized oppression exists?\(^3\)

“Internalized oppression” refers to a psychological phenomenon characterized by a person internalizing oppressive stereotypes, biases, and prejudices about an identity group to which they belong (Liebow 2016, 713). More particularly, its existence has been taken as a serious worry for the competency of oppressed people in a variety of contexts, for example: as a reason to conclude that oppressed people lack autonomy (Stoljar 2018); that they are unable to understand their own experiences (Fricker 2007); and that they lack the ability to act on their own interests (Stanley 2015).

However, feminist theorists have argued that worries about its possible effects on oppressed people are overblown or misplaced, arguing that oppressive conditions “do not typically produce pathologies or cognitive impairments; neither do the social conditions of oppression usually constitute direct coercion sufficient to erode autonomy” (Stoljar 2018).

In this paper, we advance a criticism of arguments based on the existence of internalized oppression. While critical responses against each of these competency concerns have been put forth

\(^2\) In response to claims about the epistemic advantages of the marginalized, Dror (2023, 629) rhetorically asks this question in their paper. They thank a reviewer for pushing them to say more on this point in footnote 19. We can imagine a reviewer demanding a defense of the epistemic advantage thesis during the article review process leading to the inclusion of this question and its discussion in the published work. The authors of this article have numerous examples of referees or other interlocutors offering similar criticisms to claims about the epistemic advantages of the marginalized.

\(^3\) Often this criticism comes from philosophers who are (or appear to be) dominantly situated in the field, i.e., with respect to race, gender, class, ability, sexuality, and is lodged at philosophers who are oppressed in the field across the same axes of identity and institutional position. We contend here that these two facts are related.
individually within each particular context, our approach here differs. We instead consider how some philosophers employ a general rhetorical strategy using the mere existence of internalized oppression to argue against the competency of oppressed agents, and we examine the harms this rhetorical strategy inflicts on oppressed groups. We call the use of this rhetorical strategy—wherein one moves from a claim of the mere existence of internalized oppression to the conclusion that whole oppressed groups lack competency with respect to some important kind of good—the “Internalized Oppression Worry” (IOW). To be as specific as possible, one could understand the IOW as a move from an existential claim to an at best generic and at worst universal claim of the kind: “(Some instances/at least one instance of) internalized oppression exists, therefore (generally/all) members of this particular group have diminished capacities” (e.g. “some women have internalized sexist oppression, therefore women are not capable of having autonomy”).

In this paper, we look at what arguments that rely on the IOW do and how these arguments perpetuate injustice. We argue that using the mere possibility of internalized oppression as a substantial reason to doubt the competency of oppressed agents itself does great epistemic and moral damage.

Our paper proceeds as follows. In section 1, we survey some of the ways in which the IOW has been used to argue against the competency of oppressed agents, looking specifically to the philosophical literature on autonomy and propaganda. We note the curious fact that dominantly situated people are typically the ones making these claims about how internalized oppression negatively affects oppressed people. Taking such a pattern seriously, in section 2, we consider how such a

---

4 While we discuss several philosophers who commit the IOW within the literature on autonomy, the literature on propaganda and specifically its effects on marginalized people does not have as rich of a history, and so we focus solely on Stanley’s (2015) view.
rhetorical strategy itself may constitute an injustice. There, we put forth our argument against this general rhetorical move. The IOW harms oppressed people in ways that target individuals and infect our social structures; we focus first on individual and structural epistemic harms, followed by individual and structural moral harms. Finally, in the conclusion, we explain how the IOW has real world upshots for oppressed people even outside of philosophy, and offer a series of suggestions for those considering adopting such a rhetorical move.

1. Internalized Oppression and the IOW

In this paper, we are focused on the harmful effects of the IOW’s narrative. Before that, though, we would like to briefly address the wayward concern that the IOW may be true, or otherwise a legitimate rhetorical strategy when considering the competencies of the oppressed.

First, although it is true that at least some oppressed agents do really suffer from internalized oppression, it is not at all clear how widespread of a phenomenon this is. To felicitously move from a claim that internalized oppression merely exists to the claim that whole groups of oppressed agents lack some good or capacity, one clearly needs to provide substantial empirical evidence as to its extensiveness across oppressed groups: evidence which isn’t provided in these arguments. So, the prevalence of internalized oppression is at worst overblown and at best unsubstantiated. Second, if one is concerned with agents adopting false and oppressive norms, it is common sense that the dominant are much more likely to adopt oppressive norms than the marginalized. For example, even if some women believe that their value resides in their appearance alone, many more men believe this. Liberatory epistemologists have provided a host of reasons for this phenomenon, including the fact
that such norms serve the interests of the dominant, which makes the dominant much less likely to uncover the truth of such norms (see, e.g., the rich literature on standpoint epistemology and the epistemic advantage thesis).

Finally, there are several reasons to be careful reading behaviors of the oppressed that are in accordance with oppressive norms as evidence that those agents have internalized those norms. Oppressed agents have been known to act in accordance with oppressive norms as a means of long-term resistance. One example Alison Bailey (2007) puts forth is the notion of “strategic ignorance”, when oppressed agents “play dumb” in order to gain information.\footnote{We return to this point in section 2.3.} Sally Haslanger’s (2017, 2021) work on non-cognitivist models of ideologies argues that in some contexts, “an [oppressed] individual may accept the framework [of the oppressive ideology] for the purposes of social interaction (assuming that others are ideologically shaped subjects) but not identify with their proper role in it. In such a case they would act in accordance with the ideology but not be ‘in its grip’” (2017, p. 36). Olúfẹ́mi Táíwò (2022, 44-45) similarly argues that when observing oppressed agents acting in ways that seemingly conform to oppressive norms, we should not ask why they believe these norms are true, but rather why they act as if these norms are true; he cautions us to not mistake “broad swaths of history during which oppressed populations have ‘played along’ with oppressive systems for evidence that they were ‘true believers’ in those systems”. In other words, oppressed agents may act as if they endorse certain oppressive norms for merely practical reasons: to get about the world occupied by many dominantly-situated agents who do in fact endorse such norms. Consider a woman who decides to leave the paid workforce and become a stay-at-home parent, not because she believes that she really
ought to stay home rather than have a career, but simply because society is structured in such a way that it is most cost-effective for one parent to stay home to raise children, and men are paid more than women (see Okin (1989), Cudd (2006), and Haslanger (2015)). All of these considerations give us good reason to doubt the truth of the IOW, and turn our attention to the IOW’s impact.

We also want to emphasize that we do not take issue with all discussions of internalized oppression. There is rich discourse in philosophy and critical theory that examines the phenomenon of internalized oppression to understand it better, to refine its conceptual facets and identify its siblings, to explain its features. One example is W. E. B. Du Bois describing the phenomenon of double-consciousness as seeing oneself through the eyes of another, while also seeing oneself through one’s own eyes (Du Bois 1903/2003, 9). For Du Bois, a Black man writing in the early 20th century, double-consciousness results from being a member of an oppressed racial group living in a society with a denigrating, marginalizing dominant epistemic framework, i.e., the attitudes, understandings of the world, beliefs, common sense, and concepts dominant in the social imaginary. That dominant epistemic framework is the lens through which others see him, and it shapes how he sees himself. Still, he has beliefs about himself and about Black people that are at-odds with the dominant epistemic framework. Thus, double-consciousness results: Du Bois sees himself as others do, and he sees himself through his own eyes.

Additionally, Frantz Fanon discusses internalized oppression in colonized people (1952/2008), and feminist theorists like Sandra Bartky (1990) and Robin Dillon (1997) analyze internalized oppression in women. Bartky (1990, 22) calls the phenomenon “psychological oppression”:
To be psychologically oppressed is to be weighed down in your mind; it is to have a harsh dominion exercised over your self-esteem. The psychologically oppressed become their own oppressors; they come to exercise harsh dominion over their own self-esteem. Differently put, psychological oppression can be regarded as the “internalization of intimations of inferiority”.

We take efforts like the above—efforts to understand and characterize internalized oppression—as good-faith, helpful projects mapping out conceptual space while also aiming to help oppressed people better understand their experiences. However, the IOW belongs to a second class of philosophical discussions that involve internalized oppression. These discussions employ the mere existence of internalized oppression as a logical fulcrum to argue that oppressed people are worse-off compared to non-oppressed people with respect to a variety of important moral, epistemic, and political goods, including autonomy, their ability to interrogate propaganda critically, and their reasoning and decision-making capacities. Below we survey some examples showing the widespread use of internalized oppression in the second class of discussions. These examples are in addition to the bevy of comments the authors and their colleagues in philosophy have received from interlocutors while presenting research or submitting articles for publication that argue for some advantage oppressed people have.

1.1. Autonomy

The feminist autonomy literature is a subset of the philosophical literature on autonomy which takes feminist values and concerns as its starting point. Philosophers writing in the feminist autonomy literature attempt to rehabilitate the concept of autonomy by asking, how ought we

---

6 Identity is intersectional. People can be oppressed along some dimensions of their identities while being privileged with respect to others. We think these claims are consistent with our project here.
understand the notion of autonomy—making choices based on one’s own values, desires, reasons, and intentions—in a world where those things are shaped by a patriarchal, sexist society?

Many philosophers writing about feminist autonomy reason based on both hypothetical cases and cases derived from real events and social groups. In particular, many philosophers writing about feminist autonomy employ the IOW, by arguing that since internalized oppression exists in some form, it “deforms” the choices that women make or otherwise renders those choices non-autonomous. Here are some examples:

1. Paul Benson (1991) considers a “deformed desire” case where he describes a college student who puts time and effort into practices like doing her makeup, coiffing her hair, and working out to achieve a certain body shape. Benson claims that despite being a good student with several good social relationships, the college student’s desires are “deformed” because of her internalization of oppressive norms on which her self-worth is based.⁷

2. Wendy Chapkis (1988) and Susan Bordo (1993) argue that women who undergo cosmetic surgery to more closely match socially-prescribed beauty norms are cultural “dupes” or women with false consciousness.⁸

3. Susan Moller Okin (1998) worries that some women in Other cultures are socialized into inferior submissive social roles which prevent them from critically thinking about their circumstances and realizing their authentic desires.⁹

---

⁷ One might think that because Benson is using a hypothetical case, this is an innocuous use of the IOW, since the woman in the case is specifically assumed to have internalized oppressive norms. We discuss the use of hypotheticals in the conclusion of this manuscript but briefly, using a hypothetical case like Benson does to make an inference about women generally is exactly the kind of use of the IOW that we find problematic. Benson (1991, 390) claims that the girl in the case is convinced that conforming to beauty norms is a very important goal for women in spite of her academic accomplishments, and that because she has been prevented systematically from correcting her “misunderstanding” of her own value, her motives and judgments are “less than fully her own”. But it’s exactly this kind of hypothetical case followed by a generalization that completely flattens the gendered experience of women who feel the pressure to conform to beauty norms. The case and its analysis lack the nuance required to be representative of women generally.

⁸ Bordo (1993, 260) takes up the cultural dupe notion specifically to reject the view that “talk[ing] about the grip of culture on the body” with respect to beauty norms “casts active and creative subjects as passive dupes of ideology”. For instance, many of the examples in her article focus on the racist origin of American beauty norms governing Black women’s behavior. Bordo argues that because Black beauty norms have their origin in racism, Black women who conform to these norms cannot be “creatively or playfully” (ibid, 255) engaging with them.

⁹ For example, Okin equates “parents’ giving their daughter in marriage in exchange for money” with slavery (1998, 35) and her concern is with the domestic lives of women in other cultures (ibid, 36). She claims that “respecting other cultures” has become a euphemism for denying women’s rights (ibid, 36).
4. Thomas Hill (1973, 89) discusses the “deferential housewife” who is “utterly devoted to serving her husband” and whose happiness “derives from the belief that she fulfills the role quite well”.10
5. Natalie Stoljar (2000, 95) claims that *any* “preferences influenced by oppressive norms of femininity cannot be autonomous”.
6. Anita Superson (2005, 112) argues that “many women...are at least partly duped by patriarchy”, undermining their autonomy and “full agency”.

Plenty of philosophers have done the first-order work of engaging with the above listed portrayals of women. Some reject the notion that women don’t act autonomously, or they argue that women might rationally choose non-autonomous choices, or they conclude that a women’s choice can be liberatory and authentic while still occurring within a situation where choices are limited.11 Uma Narayan (2002) characterizes European and American philosophers as having an “imperialism of the imagination” (ibid, 419-420) which causes these philosophers to Otherize women in developing countries by imagining them as either completely willing dupes of the patriarchy, wholeheartedly endorsing the cultural values by which their lives have been shaped, or otherwise “pulverized” by the patriarchy: beaten into submission with no remaining capacity for resistance. Narayan notes that women engaging in cultural practices which may seem oppressive to those outside the culture, rather than being “deformed” in some way by patriarchal cultural values, are savvy operators within certain cultural constraints; these women *bargain* with patriarchy, engaging in some cultural practices like e.g., veiling or purdah, in exchange for other social and personal benefits. This parallels Kathy Davis’s (1997, 23) claims about women who are evaluating whether to get cosmetic surgery: “women grapple actively and

---

10 Hill claims that he is talking about “servility” and he does not explicitly mention autonomy. However, as others like Andrea Westlund (2003) have pointed out, the two seem related. See also our discussion of self-respect and its relationship to autonomy on page 15 of this manuscript.
knowledgeably with opposing cultural constructions of femininity, beauty, and what should or should not be done about the female body”. Rather than being cultural dupes or coerced by patriarchal social norms, women (who appear to be acting such that they are choosing non-autonomously, based on internalized oppression) can be shrewd agents, rationally operating inside oppressive systems with critical understanding of the system’s norms. We point to the existence of these debates here as evidence that there is an inference commonly used that goes from the premise that “internalized oppression (in women) exists” to the claim that women are somehow cultural dupes of sexism and misogyny, or bad at understanding their own oppressive circumstances, or incapable of acting autonomously because they choose things that are allegedly oppressive.

1.2. Propaganda

Another topic where the phenomenon of internalized oppression has played a large role is in discussions of propaganda. In this section, we look specifically to Jason Stanley’s (2015) book *How Propaganda Works*. In it, Stanley puts forth a theory of propaganda in which it is a kind of manipulation of agents’ wills that ultimately erodes political ideals. More particularly, it does so by being a mechanism by which agents come to adopt unjust ideologies, which function to undermine the rational deliberation about one’s own and others’ interests. Given the topic of propaganda in general and the focus on how it is used to undermine political ideals, one would expect this work to be in service of undermining the oppression oppressed groups face. Unfortunately, Stanley’s claims do the opposite by taking what is fundamentally a victim-blaming stance.

The root of the issue is the way he claims the oppressed interact with oppressive ideologies. For Stanley, propaganda affects the dominant and the oppressed the same when it comes to adopting
flawed ideologies, specifically ideologies claiming the inferiority of the oppressed. According to Stanley, since propaganda’s characteristic function is to deceive people about their interests, ideals, and goals, its effect on the oppressed will be that they come to “adopt a flawed ideology of their own inferiority”\(^\text{12}\), which is just to say that they internalize their own oppression. One explanation for how Stanley envisions the oppressed come to adopt such flawed ideologies is due to the kind of political debate that takes place in the public sphere:

Subordinate group members may be led to accept, however provisionally, the negative stereotype of their group, *simply to enter smoothly into any conversation about their group with members of the dominating group* . . . Dominant group propaganda will typically propagate negative stereotypes of subordinate groups . . . It will represent members of that group as not worthy of reciprocity. So propaganda will lead to diminished self-respect on the part of the subordinate group (ibid, 163, italics in original).

Stanley also claims that subordinated groups lack alternative ideologies or narratives about themselves and their marginalization since the dominant control cultural narratives, and thus the false ideology that portrays them in negative ways is the only one they can adopt.\(^\text{13}\) What is worse, he argues that oppressed peoples’ internalization of their own oppression via adopting such flawed ideologies about themselves will ultimately prevent them from resisting their oppression and lead them to hold themselves responsible for experiencing its negative effects. Stanley (2015, 250) states, “members of the negatively privileged group will fling themselves against the high barriers erected against them, only to blame themselves for their failure to scale them,” and that, “to act against the structure that oppresses them, they need to know something about the way in which it holds them back from achievement”.

\(^{12}\) Stanley (2015, 8).

\(^{13}\) Ibid, 237.
Thus, Stanley claims that the oppressed often fail to know what holds them back from achievement, and that they falsely adopt oppressive ideologies.

There are several issues with Stanley’s understanding of the inner mental lives and actions of oppressed groups. First, there is the claim that what prevents the oppressed from successfully resisting oppression is their own beliefs about their inferiority rather than social-structural forces like the constant threat of violence, and that, in the rare cases the oppressed do resist, they blame themselves rather than understanding their failures to scale “the high barriers erected against them” are further evidence of the deep injustice of the society in which they live. Second, there is the claim that the oppressed do not have alternative subversive narratives of their own construction. However, such narratives largely remain within their own communities due to the willful ignorance of dominant groups (see Pohlhaus (2012) on willful ignorance; Mason (2011) also argues that oppressed communities have much greater narrative and conceptual resources than the dominant suppose).14 Lastly, we can discern the root of these misunderstandings of the lives of oppressed groups to be his claims that the oppressed know very few facts about their own oppression or “what holds them back from achievement”, rather than being the most aware out of all socially positioned agents on these matters, and therefore the least likely of all to actually internalize such ideologies. Veronica Ivy (publishing as Rachel McKinnon) in her (2018) review of How Propaganda Works voices such a criticism when she argues that the long-established literature on standpoint epistemology shows “why negatively privileged persons...do not need to be ‘exposed’ to an alternative ideology to combat the

---

14 Here one should note that although Stanley is marginalized across ethnicity as a Jewish man, he is dominantly situated in largely all other respects—notably gender, class, and race, which afford him significant power—while predominantly interrogating from the outside groups that lack such power.
propaganda”, as oppressed people tend to know that they’re oppressed, and they have accurate, detailed understandings of the sources of their marginalization. Given this, we have substantial reason to believe that Stanley’s baseline assumption—that the oppressed are generally just as likely to adopt false ideologies as the dominant—is false.\(^\text{15}\)

Moreover, there are further issues to discern when one takes a step back from these claims themselves and considers the bigger question of what such claims do. How does advancing such a picture of the inner mental lives of the oppressed affect the oppressed themselves or the existence of oppressive ideologies? Gaile Pohlhaus (2016, 354) briefly raises such a concern when she states, “I am worried about what these arguments do. ... [They] ... encourage us to think of non-dominantly situated persons as causally related to one another in a manner that does not engage them as full epistemic agents.” In the next section, we take just this approach, putting forth an analysis of the many ways in which arguments that use the mere existence of internalized oppression to argue that the oppressed are incompetent in some way harm the oppressed and perpetuate their circumstances of oppression.

2. The Harms Arguments Based on the Internalized Oppression Worry Inflict

\(^{15}\) No doubt there are cases in which oppressed people do believe false narratives about their own oppression. For example, one might point to the fact that white, American women often support patriarchal ideologies associated with right-wing political beliefs or some sects of Christianity. Again, we are not denying that internalized oppression exists, or that some oppressed people full-throatedly endorse the ideologies that oppress them. But far more often, oppressed people have nuanced and complicated reasons for engaging with or supporting oppressive ideologies. For example, white women also benefit from white supremacy, which is also part of the aforementioned right-wing ideologies. So, even a full-throated endorsement of that ideology is complicated given that the ideology is simultaneously beneficial and oppressive.
In this section, we analyze the harmful effects of arguments that use the existence of internalized oppression to argue against the competency of oppressed groups, focusing on two arenas of harm and two parameters within these arenas: epistemic and moral harms, and, within these, harms to the individual and structural harms. Although these harms are often experienced simultaneously and are interlocking, we conceptually separate them for the sake of clarity. Within the epistemic harms incurred to the individual, we argue that the IOW infantilizes the oppressed, and attempts to undermine their epistemic self-trust; regarding epistemic structural harms, we argue that the IOW functions to reproduce oppressive epistemic hierarchies, and allows social elites to monopolize the epistemic good of social attention to reinforce existing power structures. We continue that the IOW is morally bad in its patronizing denigration of the faculties of the oppressed. This denigration damages individuals’ moral agency. Furthermore, the IOW creates unjust moral hierarchies by constructing false moral narratives of both the characters of the oppressed and what ought to be done given facts of oppression.

2.1. Individual Epistemic Harms

Arguments that employ the IOW negatively portray the capacities of the oppressed to make accurate judgments about themselves and the world. The autonomy literature claims that the oppressed are unable to make accurate judgments about their own desires and values, and what is good for them; the propaganda literature claims that the oppressed are unable to form accurate beliefs about why they fail to achieve, the nature of their own characters, and the nature of oppressive structures themselves. As we have already noted, there are serious reasons to doubt that any of these claims about the oppressed are true. However, these claims attempt to impair oppressed peoples’ epistemic
autonomy. Claims that the oppressed cannot form accurate judgments about themselves and the world can create cognitive tension in the oppressed, making decisions about oneself harder and making it harder to understand one’s own identity. Therefore, claims based on the IOW harm the oppressed by creating unnecessary epistemic obstacles and cause suffering in some cases. However, oppressed people can still typically make fulfilling decisions about their lives and come to understand their own identities.

Consider that to have epistemic autonomy—autonomy in forming beliefs, in assessing evidence, in discerning true experts, etc.—one must have a healthy degree of self-trust. This is because epistemic autonomy requires one to operate independently from others when performing crucial epistemic tasks like those noted above. If one lacks self-trust, one lacks reason to be epistemically self-reliant, as one believes that one cannot reliably form accurate beliefs and judgments. We contend that narratives of the mental incompetency of the oppressed attempt to erode the self-trust of members of oppressed groups. Using the fact that internalized oppression exists and affects members of oppressed groups works as a kind of gaslighting: a reason for the oppressed to doubt their own

---

16 We think there are interesting and important questions about the relationship between academic philosophy and the social imaginary. How are narratives of the mental incompetency of the oppressed communicated to and within the social imaginary? Does academic philosophy discourse get picked up by mainstream discourse, or does academic philosophy take up notions from the social imaginary and develop argumentation for them? Although we do not have the space to explore these questions fully here, we can offer some comments. In particular, it seems like academic philosophy does have some effect on the social imaginary through venues like mainstream media and social media, especially when academic philosophers are commenting on hot-button social and political issues. The New York Times reviewed Jason Stanley’s 2018 book How Fascism Works, which contains some of Stanley’s ideas on propaganda. Other philosophers are regularly featured in the Times’ The Stone column. Many contemporary philosophers have a social media presence as well. For instance, Kathleen Stock is a philosopher who has extensively argued against trans rights. As of the time of writing, she has over 146,000 Twitter/X followers. Her social media postings on her alleged censorship due to her anti-trans work were cited in a US Supreme Court Amicus Brief (Carter 2019); the brief was written in support of the position that employers could discriminate against trans people on the basis of gender identity. So, it does seem academic philosophy’s discourse permeates the social imaginary through various entry points in mainstream media and social media. It is not hard to imagine that narratives about the mental incompetency of the oppressed, i.e., the IOW, in academic philosophy can therefore enter the social imaginary through these points as well.
judgments. Crucially, the more these internalized oppression worries are raised, and the more people who raise them, the more one’s self-trust may be affected, since it is primarily through the feedback of others that self-trust is formed (Jones 2012). We use others’ judgments as confirmation of the accuracy of our own, and our overall epistemic competency: the more others affirm our judgments and assess us as credible knowers, the more we trust in our own epistemic abilities, and the more others disconfirm our judgments and assess us as incompetent failed knowers, the more we doubt our own epistemic abilities.

Undermining an oppressed person’s self-trust, an epistemic harm, can compound into a moral harm. Joel Anderson and Axel Honneth (2005, 130-131) hold that autonomy involves things like being able to sustain attitudes of self-trust, self-respect, and self-esteem. According to Anderson and Honneth, these properties emerge as the result of interacting with other people, and seeing that others value oneself as an object of concern, a responsible agent, or a contributor to shared projects. Similarly, Benson (2005, 108) offers an account of autonomy on which autonomy involves “taking ownership” of one’s actions and intentions, meaning one is in a position to speak on behalf of one’s actions and intentions. In order to take ownership of one’s actions and intentions, one must trust oneself as being capable of making informed and justified decisions. As we can see on both Anderson and Honneth and Benson’s account of autonomy, self-trust is required for autonomy, and self-trust has itself a dependence on social relations. So, if the IOW undermines self-trust, that poses a threat to autonomy.

Additionally, the harm of attempting to undermine an agent’s self-trust carries with it real material harms, especially for oppressed people. Consider the fact that trans people have historically
been “gatekept” from gender-affirming medical treatment.\(^{17}\) “Gatekeeping” refers to the phenomenon by which trans people seeking gender-affirming care need to “prove” that they are trans and therefore treatment is “medically necessary” to receive, e.g., surgical interventions, hormone replacement therapy, and other forms of care. Part of this proof involves trans people demonstrating a “marked and sustained” gender incongruence and that “other possible causes of apparent gender incongruence have been identified and excluded” according to current World Professional Association for Transgender Health (WPATH) guidelines (Coleman et. al, Appendix D). Demonstrating a stable and uncomplicated gender identity absent other signs of neurodiversity has historically been used to gatekeep trans people from gender-affirming care as well.\(^{18}\) Because of the history of gatekeeping, trans people and people questioning their gender identity are highly sensitive to worries about being denied medical treatment. Consider a comment from the Reddit subforum r/asktransgender, in a post titled “self-doubts regarding medical treatment and autism”:

I’m scared of bringing [my gender-identity doubts] up with a doctor because I feel like it could have an adverse effect on my treatment that I’ve been in the queue for for [sic] the better part of 2 years now. I’ve gone through people prodding about my past, and I also went to a “full” psychological evaluation a few months back but my results from that have not come back. On the other hand I’m also terribly afraid that this is simply some obsession that might go away if I would dig deeper into my neurodiversity, because I’ve noticed I obsess over lots of things for varied amounts of times. I also understand there are a very large amount of people who are autistic and transgender and treatment was absolutely the perfect solution for them (“self-doubts regarding medical treatment and autism”, 2021).

The commenter has two worries: that they might be wrong about their own identity, and that they are worried that discussing this with a medical professional will impair their ability to access treatment.

\(^{17}\) See Ashley (2019).

Having self-doubts is a normal part of considering a major life change; one might worry about whether they really ought to buy a house or continue renting, whether having children is a good decision, and similarly, one might worry about whether receiving treatment like hormone replacement therapy is what one ought to do. But trusting oneself means trusting one can adjudicate among one’s worries and make a good decision. In the case of the Reddit commenter, a lack of self-trust might ultimately harm the commenter; if they choose to bring up their worries with a medical provider, they might be gatekept from accessing medical care. If they receive gender-affirming care without adjudicating their doubts, they might ultimately regret that decision. Despite cases like the above, many people questioning their gender identity that are able to navigate the medical system and receive gender-affirming care make good, fulfilling decisions. Regret following gender-affirming medical interventions are rare (Bustos et. al 2021, Turban et. al 2021). This is just one type of example of oppressed people having epistemic autonomy and employing their cognitive capacities to make robustly good decisions for themselves, contrary to the way oppressed people are represented by the IOW.

In sum, arguments that attempt to erode the self-trust of the oppressed fail to do that. Oppressed people are still able to act with epistemic autonomy as they consider themselves and the world. But arguments that employ the IOW (like Stanley’s, or Okin’s, or others) and gaslight oppressed people about their capacities make their epistemic, and, consequently, practical lives harder. These arguments make their decisions harder than they ought to be by contributing to a social imaginary portraying them as cognitively and epistemically incompetent. This is harmful and unjust, and can contribute to material obstacles to the well-being of the oppressed.
2.2. Structural Epistemic Harms

We can understand epistemic hierarchies as ways in which social groups and individuals (in virtue of their membership within social groups) are structured with respect to epistemic goods like credibility, trust, expertise, knowledge, etc. These epistemic goods bear crucial relations to social and political power, so these goods are one way in which social-political hierarchies are reproduced and enacted. We can understand this by looking to the widely discussed case of credibility judgments. Credibility judgments—judgments about the trustworthiness of agents and their knowledge with regards to specific subject matters—are significantly affected by the social status and identity of the agents themselves, such that those with oppressed status are generally judged as less credible than those dominantly situated (Fricker 2007); this is how we can see one’s place within the broad social-political hierarchy affecting one’s position within the epistemic hierarchy. However, the way these hierarchies interact is also mutually supportive and reinforcing, as one’s place within the epistemic hierarchy can reinforce one’s place within the social-political hierarchy (see Dular 2017). Taking our case of credibility, we can see that being judged as not credible and lacking in knowledge can reinforce and serve as a justification for one’s low status on the social-political hierarchy, e.g., that one ought not occupy positions of power that require expertise like certain career roles of authority, prestige, and financial power. For instance, consider a woman of color in a corporate environment; it’s unlikely she will become CEO and gain material power if she is judged as lacking credibility. It is our claim here that the IOW, when used as an argumentative strategy to claim the oppressed are incompetent, functions to solidify these unjust epistemic hierarchies.
Since the dominant are much more likely to adopt false ideologies than the oppressed, the oppressed are the true experts on these matters, and epistemically ought to be in superior positions within the epistemic hierarchy. Unfortunately, this role reversal of expertise is not unique to the IOW. Rather, it’s one specific way in which the dominant strategically force the oppressed into lower positions of epistemic power and themselves in positions of expertise. For example, Dular (2021) argues that mansplaining should be understood as a forceful subversion of the epistemic roles of hearer and speaker in a testimonial exchange wherein dominantly situated agents (specifically men) make a power grab with respect to expertise. Linda Alcoff (1991-1992) also notes this in her piece “The Problem of Speaking for Others” when she states:

I would stress that the practice of speaking for others is often born of a desire for mastery, to privilege oneself as the one who more correctly understands the truth about another’s situation or as one who can champion a just cause and thus achieve glory and praise. And the effect of the practice of speaking for others is often, though not always, erasure and a reinscription of sexual, national, and other kinds of hierarchies (ibid., 29).19

Our contention here is that the IOW also institutes this role reversal of expertise and is used to reinforce oppressive epistemic hierarchies, functioning to give the dominant more epistemic power with respect to matters of oppressive ideologies than they are warranted. Moreover, and what is worse, the dominant can now justify their superior position of epistemic power, as the IOW provides a narrative to which they can refer about why the oppressed should not be trusted nor thought of as relative experts on matters relating to their own oppression. This epistemic justificatory narrative crucially relates to similar moral narratives discussed in section 2.4 below.

19 Sometimes speaking for others is a pragmatic tactic for the aims of social justice; see (Lance 2021). These are not the kinds of ‘speaking for’ cases with which we are concerned, though.
There are other structural epistemic harms caused by the IOW, viz., the epistemic good of attention. Cultural conversation is structured by elites; people who are dominant in terms of cultural cache, social and economic class, and the institutionally and politically connected determine what is on the front page of the *New York Times* every day and how the news is framed (Táíwò 2022). Thus, cultural and social attention is controlled by these elites. Since already dominant groups and people control cultural conversation and media narratives, they determine things like the content of the social imaginary, as well as political priorities and initiatives. This harms oppressed people in an epistemically significant way by diverting attention from the conditions of their oppression to alleged internalized oppression.

For another example involving trans-related concerns, recall that there is a prevailing medical narrative that people who are questioning their gender identity might be confused because of neurodiversity. There are other cultural narratives about the validity of trans peoples’ identities, for example: that people who would formerly have been gender non-conforming in some way, like “tomboys” and butch lesbians, are having their preferences related to behavior and presentation medicalized and thereby coerced into transitioning.20 This narrative says something like: “when women who would be tomboys internalize misogyny, they learn to hate themselves as women. And transitioning into a man solves that problem. But they didn’t need medical intervention. Instead, they needed to accept themselves for who they really are and learn to love being a woman.” So, the narrative claims that women are getting confused about their gender identity because of internalized misogyny.

---

20 Janice Raymond’s (1979) asserts this myth but there are lots of other recent examples too; see Jeffreys (2014), Moore and Brunskell-Evans (2019), and Soh (2020).
When mainstream publications like the *Wall Street Journal* publish concerned editorials lamenting “transgender ideology” and its excise influence on gender nonconforming people who otherwise wouldn’t want to transition, that diverts cultural attention away from actual problems that gender-nonconforming people face. Trans and gender-nonconforming individuals are subject to discrimination and violence in almost every social arena of life, including school, employment, and medical care. Besides diverting our cultural attention, these kinds of narratives are often used to justify oppression: the notion that innocent children need to be protected from “transgender ideology” is currently galvanizing an anti-trans political movement across the United States.

What should we be more concerned about if we want to live in a just world? Attention is a finite good, and there is only so much space in the *Wall Street Journal*. Should we be focused on the possibility that a small number of people who transition—1%—regret it? Or should we be focused on the physical and psychic harms suffered by the other 99%? Should we be focused on the possibility that a small number of oppressed people have come to endorse oppressive norms? Or should we be focused on the ways the world is structured to oppress them? The IOW causes structural epistemic harm to oppressed people, which can have material effects.

### 2.3. Individual Moral Harms

Beyond epistemic harms, the IOW also generates a multitude of moral harms, directly denigrating oppressed agents and undermining their ability to function as equals in the moral community. Even if it is raised in good faith, the IOW impugns the capacities of the oppressed: their

---

21 Wright 2022.
22 James et al. 2016.
23 Tanner 2023.
capacities for autonomy, their epistemic faculties, their critical interrogative faculties. This is not a normatively neutral act, but rather denigrates oppressed agents. We submit that the denigration involved in the IOW also corrupts individuals’ ability to function as full moral agents.

In her 2016 paper, Nabina Liebow argues that oppressed agents who genuinely do come to internalize oppression may have damaged moral self-conceptions wherein they view themselves as moral deviants or as having a reduced sense of moral agency. Due to these skewed self-perceptions, oppressed agents might fail to feel that they can be held morally responsible or feel that they are morally deficient and unworthy. These feelings, Liebow argues, lead to being ostracized with respect to the moral community via being placed outside it, or within it but as antithetical to it. We think the IOW can likewise corrupt an individual’s ability to function as a full moral agent: not in virtue of how those individuals view themselves, but in virtue of how others in the moral community who advance the IOW and believe conclusions derived from it view them.

To participate as a full moral agent in a moral community an individual must not only have certain features or competencies, but also must be viewed and treated by others in a certain way. For example, one must be capable of controlling one’s actions and holding oneself accountable, but others must also hold one accountable for one’s actions. Liebow claims internalized oppression causes others to view oppressed agents as incapable or criminal, which causes oppressed agents themselves to adopt these narratives as their self-conceptions and place themselves outside of the moral community. In the case of the IOW as a rhetorical strategy, we submit that the chain of events works in an alternative manner: namely, others view oppressed agents as incapable, which causes these same others to refuse to let oppressed agents participate as full moral equals in the moral community.
What the IOW does is move from a claim about the mere existence of internalized oppression to the conclusion that oppressed agents are incompetent, especially with respect to discerning and acting on behalf of their own interests and desires, i.e., their autonomy. Having these capacities is crucial for functioning as a competent moral agent in a moral community. But it is also crucial that others in the moral community believe that one has these capacities.

It is in precisely this way that the denigration we contend is at the heart of the IOW undermines the ability for oppressed agents to function as full moral agents in the moral community. It’s not in virtue of the way oppressed people in fact are or come to see themselves, but rather in the way that others come to view and treat them. Because of beliefs formed on the basis of the IOW, at best others may come to view oppressed agents as akin to children: those with moral rights but diminished moral agency, who need to be controlled rather than engaged with as an equal.

Moreover, the belief in the incompetency of oppressed agents morally harms them by portraying them as both committing more morally wrong actions, and less morally good ones, than they are likely committing. This, too, undermines their status in the moral community. First, the IOW concludes that oppressed agents act against their own interests. In this way, oppressed agents can be viewed as wronging themselves via failing to live up to their obligation of self-respect and general Kantian obligations to the self. Second, by concluding that the oppressed act against their own interests, the IOW implies that oppressed agents fail to resist systems of oppression, which is itself a morally good action. This encourages a misperception of the actions of the oppressed that are indeed

---

24 Some argue that the oppressed have a moral obligation to engage in acts of resistance against their own oppression; see Hay (2011).
acts of resistance, as instead failures to resist. As mentioned in section 1, this misperception has been thoroughly discussed by Alison Bailey (2007), where she conceptualizes “strategic ignorance” as the ways in which the oppressed knowingly rely on the dominant’s degrading misperceptions of them to better resist their own oppression. For example, Bailey discusses the case of Frederick Douglas, who relied on whites’ misconception of Black people as incapable of being literate or clever to get free help on his political writings of resistance, as well as enslaved people relying on whites’ misconception of Black people as clumsy and bumbling to mask their intentional actions of retribution like breaking property, and therefore avoid penalty. In this way, the implications of the IOW promote the erasure of oppressed agents engaging in morally good acts (of resistance)— actions crucial for being perceived as an equal in the moral community—thereby undermining their ability to operate as a full moral agent in the moral community.

2.4. Structural Moral Harms

We believe that conceiving of the moral realm in a structural way, as related to structures and practices of oppression, can helpfully illuminate some further issues with the IOW and how it harms.

We can think of moral structures as those that structure the moral landscape of the world, which itself is composed of actions and characters, broadly. Moral structures then divide the moral landscape of actions into categories of good, bad, morally obligatory, supererogatory, etc.; they divide the moral landscape of characters likewise into those that are good and bad (i.e., the virtuous or moral heroes, and the vicious or moral villains), or those characters that lead one to (reliably) perform morally good actions or bad ones (e.g., analyses of the kinds of motivations that are morally best). We can conceive of moral narratives as an explanatory tool within the construction of the moral landscape that
explain how and why such actions and characters are hierarchically sorted the way that they are, as narratives broadly tell us the kind of actions to do and not to do, the kind of people to be, and why. However, these moral narratives, like any other kind of social narrative, are shaped by the actual cultural milieu of the society in which we live. This means that, despite these narratives being moral in nature, they are not immune to being tainted by the actual oppressive ideologies that exist. In this way, moral narratives can themselves be oppressive, as they can structure the moral landscape in ways that ultimately solidify oppressive hierarchies and harm oppressed groups.

The IOW, when used to argue for the incompetence of the oppressed, functions as one type of toxic, oppressive moral narrative. We argue here that it puts forth a false understanding of what morally ought to be done in contexts of oppression, as well as a false story about the moral characters of the oppressed and the dominant.

First, consider how arguments employing the IOW may generate an understanding of what morally ought to be done about oppression. As we have mentioned previously, the IOW claims that the oppressed misunderstand non-normative facts about their own oppression, and are therefore in an epistemically bad position to make correct normative judgments about what ought to be done regarding their own oppression. We then might ask, who is better positioned to make such normative judgments? People without internalized oppression, i.e., the dominantly-situated. Recall that the IOW causes epistemic structural harms, raising the credibility and epistemic expertise of those dominantly situated. Also notice that the narrative constructed using the IOW contains within it the assumption that those dominantly situated know what is best for the well-being of the oppressed, since that narrative claims that the oppressed regularly act against their own interests due to internalized
oppression. So, the moral narrative constructed is that those dominantly situated ought to help or fix the oppressed, and to decide what ought to be done when it comes to circumstances of oppression itself.\textsuperscript{25}

One way that we can understand the IOW is as part of a racket of domination. Writing on rackets operating within contexts of oppression, Claudia Card (1991) argues that rape can be understood as a racket operating within the context of patriarchal society. She (1991, 304) defines a racket as those practices or institutions that “\textit{create} danger to sell ‘protection’”. In the case of rape, the threat of rape and its prevalence make it so that women must seek out male ‘protectors’. If they don’t, women are held responsible for what happens to them; they only have themselves to blame if they are sexually assaulted. However, it is men who create the danger and constant threat of rape for women. Thus, men create the danger of rape and then present themselves as the protection women need from that exact danger. We contend that the IOW constructs an oppressive racket via serving one aspect within the racket of domination. It is the dominantly situated that are responsible for both the original construction of oppressive ideologies and practices, and the dominant who largely maintain their existence (Frye, 1983, 12-13); this includes the construction of oppressive norms, values, and stereotypes that the IOW claims the marginalized adopt. This is the danger that is created. With the moral narrative that the dominant are best situated to determine what is best for the well-being of the oppressed and what ought to be done within circumstances of oppression, we get the second aspect of

\textsuperscript{25} Importantly, one might think that it’s true that the dominant are obligated to help the oppressed or fix issues of oppression, as they owe the oppressed this labor as one means of creating equity given the unarmed advantages they reap from the existence of others’ oppression. We take no issue with this understanding of what the dominant owe to the oppressed and why. The issue we take here is with the claim that the basis for the claim that the dominant ought to help is that they are in an epistemically superior position. We are grateful to an anonymous reviewer for making this matter clear to us.
the racket: the dominant presenting themselves as the “protection” or solution to the danger that they themselves have created.

The story being told with the IOW—that the oppressed need help from the dominant—constructs a picture of the oppressed as the villains in their own story, and establishes the dominant in a savior role. In this picture, the oppressed are the ones who enact and maintain oppressive ideologies, failing to resist their own oppression and believing the noxious stereotypes about themselves; thus, the oppressed are categorized as morally bad, as they are the ones responsible for the harm they suffer. On the other hand, the dominant, being more competent with respect to accurately discerning the nature of oppressive ideologies and making judgments concerning well-being, are portrayed as moral innocents in an unjust world, there to save the oppressed from themselves. This story, then, is very similar to the white savior complex widely discussed, wherein white people view themselves as the ones who are to save communities of color from racism and colonialism. Really, though, since it is white people who create and maintain the racism and colonialism to begin with, they are no solution and no savior at all. A similar narrative is criticized by Marx in his analysis of bourgeoisie-proletariat relations. The moral narrative involving the characters of the oppressed and the dominant is backwards, as the dominant are the real villains and the oppressed their own saviors.

3. Conclusion

In this paper we have argued that the general rhetorical move by some philosophers to reason from the existence of internalized oppression to the impaired or deficient capacities of the oppressed is epistemically and morally harmful to oppressed people. We think that intuitively, these epistemic and
moral harms provide *pro tanto, prima facie* reasons not to employ the IOW. The IOW’s harmfulness also gives us metaphilosophical reasons to think about the rhetorical effects of arguments employing the IOW. What do those arguments contribute to philosophical discourse? Are arguments that denigrate and harm oppressed people worthwhile philosophical pursuits?

One way to think about the answers to these questions is by asking what we gain from arguments based on the IOW. By examining these arguments, do we gain some illuminating philosophical truth? Arguably, in the examples we have showcased, that answer is “no”. It is surely true that, for example, some women endorse the patriarchal system that oppresses them, and we can imagine women who confabulate reasons for undergoing cosmetic surgery that hide the bending of a woman’s will by coercive beauty norms. But as we have seen, women are a diverse group of people. Discussions of their autonomy and the reasons why they make certain decisions in light of an oppressive system require a nuance that reflects that diversity. Arguments that generalize from the mere existence of internalized oppression to some feature of “women” simpliciter (or even “American women” or “Korean women” or “Sufi women”) are bound to fail because they overgeneralize. Perhaps we gain a deeper understanding of a notion like autonomy from these discussions but it is difficult to measure philosophical progress against the real harm that these discussions inflict.

We urge philosophers, and other theorists, who theorize about the lives of oppressed people to begin with charitable assumptions: when it seems as though oppressed people are engaging in acts against their own interests or in line with oppressive norms, one ought to consider that there is a very good reason why they are doing so in a way that is consistent with their own interests, rather than concluding that the oppressed are confused or deluded and thus incapable of critical thought or
autonomous action. We also urge philosophers who theorize about the lives of oppressed people to critically consider their reasons for doing so. Remember here Alcoff’s (1991) claim that speaking for others is often born out of a desire for praise and social esteem, to be seen as a moral champion and enact mastery over others. This is especially important given the current climate of academic philosophy, where it has become more fashionable to contribute work on matters of oppression and injustice now that this kind of work is no longer kept at the absolute margins of what’s considered “real philosophy”; consider, for instance, the recent mainstream philosophical interest in social ontology. If one has only recently entered into a discussion about the capacities of the oppressed, consider whether one is contributing to a conversation in a novel way; we think there is a good chance philosophers working on liberatory epistemologies and moral theories are not hearing the IOW for the first time.

Finally, we urge philosophers to employ intersectional analyses to understand their own identities with respect to the groups of oppressed people that they are discussing. One ought to consider one’s position as a member of more dominant groups (white, cisgender, non-disabled, wealthy, tenured) when advancing the IOW against oppressed groups and other philosophers’ theories. Philosophical discourse does not happen in a vacuum, and claims about the lives of the oppressed are not normatively neutral. In her meditation on contemporary discussions about trans people in philosophy, Talia Bettcher offers this reflection:

I’m afraid there’s a tendency among some philosophers to suppose that philosophical investigations into race, gender, disability, trans issues, and so forth are no different methodologically from investigations into the question whether tables really exist. One difference, however, is that while tables aren’t part of the philosophical conversation, trans people, disabled people, people of color, are part of the conversation. Or at least, we think we are. We’re here. In the room. And we’ve suffered from life-long abuse. I’ve helped a friend die
of AIDS, fending off the nurses who misgendered her, watching in horror as the priest invalidated her entire life at her funeral by reducing her to a man. I’ve been personally assaulted in public to prove that I was a man. I’ve had a friend trans-bashed. And as this beating was gang-related, she then lost her home. I’ve had a friend stripped by police-officers, forced to parade back and forth while they ridiculed and harassed her. So please understand that this is a little bit personal (Bettcher, 2018).

Bettcher reminds us that philosophizing about oppressed peoples’ lives means philosophizing about actual terrible things that happen to people living under unjust conditions.

We can see that the way that the IOW is used in the discussions of autonomy and propaganda specifically carries real world material harms. Whether we think that women of color who veil can make autonomous choices about their lives matters to legal statutes about whether they are allowed to do so in public. Whether we think that poor people of color choosing to not vote in democratic elections is them acting against their own interests matters to whether they are accepted as part of our civic body. In light of the negative consequences that arguments based on the IOW can generate, then, we must take seriously the notion that those arguments ought not be made, in order to prevent harm.

We emphasize that we believe that theorists employing the IOW philosophize in good faith and ask important questions. Jason Stanley’s How Propaganda Works asks why ostensibly oppressed people often vote against their own interests. The conclusion of our paper is not to stop asking important questions. Nor do we insist on purity of identity when it comes to philosophical theorizing. Our conclusion is that philosophers must reconsider making arguments that ask questions like Stanley’s and then conclude that the answer has to do with internalized oppression that makes oppressed people epistemically, morally, politically, or cognitively deficient. Philosophers need to be charitable when it
comes to theorizing about the capacities of the oppressed and they must engage with testimony regarding the lived experience of having an oppressed identity, to do those identities justice.
References


---


Tanner, Lindsay. (2023, March 5) *How Common is transgender treatment regret, detransitioning?* AP News.


Author Biographies:

Nicole Dular is an Assistant Professor of Philosophy at Notre Dame of Maryland University. She works on issues at the intersection of ethics, epistemology, and feminist philosophy. Her work has been published in *Philosophical Studies, Hypatia*, and *Feminist Philosophy Quarterly*, among other peer review journals. You can find more about her research and teaching at nicoledular.weebly.com.

Madeline Ward is an Assistant Professor of Philosophy at Western New England University. Her research focuses on issues in feminist epistemology, anti-oppression philosophy, and bioethics. More of her work, including public philosophy, is featured at madelineward.com.