Hypatia (2026), 1–22
doi:10.1017/hyp.2026.10081

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Sensitivity and Survival

Kate Manne

Sage School of Philosophy, Cornell University, USA
Email: kate.manne@gmail.com

(Received 29 July 2024; revised 8 October 2025; accepted 2 March 2026)

Abstract

Accusations of over-sensitivity are common, particularly when it comes to forms of systemic oppression such as racism and misogyny. I consider three things such over-sensitivity could plausibly be: a disposition to over-identify instances of these phenomena, the over-expansion of the relevant concepts, and a tendency to overreact to putative instances thereof. In each case, I find reasons why such worries are likely to be overblown. In the first, an agent's supposed over-sensitivity to racism and misogyny is often simply sensitivity to phenomena which are routinely minimized; in other cases, her slight over-sensitivity nevertheless has value in a social environment in which under-sensitivity is a large and predictable problem. In the second case, the supposed over-expansion of the relevant concepts often corrects for their artificial narrowing, and makes us more sensitive to genuine harms and wrongs. In the third case, an agent's supposed overreactions to racism and misogyny reflects an affective disposition that constitutes a valuable defense against the twin perils of gaslighting and numbness. I conclude by considering the possibility that there is a real strand of over-sensitivity to harms to dominant group members in the current discourse that illicitly serves their interests and functions to silence people facing oppression.

When the Indian-American comedian, Hari Kondabolu, was performing onstage at the Edinburgh fringe festival in 2011, he had a familiar experience: being heckled by an audience member. He angrily called it out. But it turns out he was mistaken—the yelling hadn't been heckling to the effect that he was a Palestinian terrorist (by saying "Palestinian power!" sarcastically, as he'd thought). In fact, the yelling was not a racist remark of any kind. Rather, the supposed heckler had actually been echoing Kondabolu (by saying "Men are in power!" in answer to his question of why we've never had a female president). Even after Kondabolu duly apologized for his mistake—and his anger—an audience member accused him of being over-sensitive. "No, just sensitive,"

© The Author(s), 2026. Published by Cambridge University Press on behalf of Hypatia Inc. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

he responded, going on to recount the incident on the popular storytelling radio show and podcast, *The Moth*.[1]

Accusations of over-sensitivity are nowadays very common. Yet, with notable exceptions, little attention has been paid to this kind of criticism in the feminist philosophical literature to date.[2] This paper aims to correct this omission. I will proceed from the assumption that being *sensitive* to some morally or politically significant property in the world—say, racism—is a disposition of an agent, and is evidenced over time, rather than being a trait we can say she has on the basis of just one occasion. And I take it that this disposition involves a certain pattern of responsiveness: detecting true positives and true negatives, and avoiding false positives and false negatives, with respect to the target property, with sufficient reliability. (I leave it open what this threshold for "sufficiency" is, exactly.) It will also involve having appropriate (or, similarly, apt or fitting) affective reactions to the instantiation of said property—feeling disapproval, dismay, disavowal, and concern about racism, for example.[3] Having understood sensitivity in this way, there are three natural things for over-sensitivity to be: first, being over-sensitive might involve a disposition to over-identify *instances* of the phenomenon in question, for example, sounding false alarm bells for racism where none is actually present. Second, being over-sensitive might involve a disposition to use concepts that are overly extensive, for example, in operating with a concept of racism that is implausibly expansive. Third, being over-sensitive might involve being prone to overreact, affectively, to the putative truth of the relevant ascription, for example, by being gripped with rage when more modest consternation is called for in response to some small and isolated incident.[4] I will argue in this paper that worries of all three of these kinds of over-sensitivity with respect to racism, misogyny, and other, similar forms of structural oppression are typically overblown. Indeed, the shoe may be on the other foot: the over-sensitivity of dominant group members may be distorting the discourse, and in ways that are not innocent.

This paper can be understood as a contribution to two rich literatures: feminist epistemology and non-ideal theory. On the former score, feminist epistemologists following Helen Longino (1990), Patricia Hill Collins (1990), Sandra Harding (1991), Elizabeth Anderson (1995), Louise Antony (2000), and many others have long emphasized the fact that knowledge and beliefs are acquired within *communities*, and that even scientific inquiry must be examined empirically as a practice that actual human beings—with our biases, limitations, values, and particular social position—are imperfectly engaged in. Relatedly, non-ideal theorists have drawn valuable attention to the way that lessons drawn from theorizing as if the world is epistemically frictionless

[1]Hari Kondabolu, "Not TOO sensitive, just sensitive," *The Moth*, July 1, 2014. https://player.themoth.org/#/?actionType=ADD_AND_PLAY&storyId=737.

[2]See, for example, José Medina's 2012 treatment of sensitivity and what he calls "sensitivity to insensitivity," for an orthogonal but related discussion of a critical awareness of others' epistemic injustices. See also Taylor Rogers's 2021 treatment of the role of affective numbness in promoting what Kristie Dotson calls epistemological resilience, where an epistemological system remains stable in the face of counter-evidence. In this paper, I will focus on affective numbness primarily as a response to oppression rather than playing an important role in perpetuating it, per Rogers, but I agree that the phenomena they point to are real and important.

[3]Must the reaction be proportionate too? I'll come back to this issue in the third section of this paper.

[4]We can say the same, *mutatis mutandis*, for under-sensitivity. My thanks to Quill R Kukla for discussion here.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

may not generalize very well to our messy, real-world context. To take the metaphor offered by Charles Mills, in his classic paper "Ideal theory as ideology":

> A very smooth, Teflon-coated plane suspended in a vacuum may come close enough that one can regard its behavior as approaching that of an ideal frictionless plane … But if the plane is covered not with Teflon, but Velcro, or is pitted, cracked, and abraded in various ways, then obviously this would be absurd. (Mills 2005, 167).

The abrasions I am interested in here, to extend Mills's metaphor, derive from the fact that people are not, all too obviously, perfectly attuned to many important social phenomena: racism, sexism, and misogyny included.[5] While the ideal would be for us all to be perfectly sensitive—and thus neither over-sensitive nor under-sensitive, and indeed able to get it right in every conceivable circumstance—the fact that different agents predictably fall short of this ideal in different, and patterned, ways makes an important difference to how we should criticize ourselves and each other. It also makes a difference to how we should view small failures and foibles in the vicinity of over-sensitivity: we shouldn't overreact to them. Or so I will be arguing.

1. Over-identification of instances

The first kind of over-sensitivity involves over-identifying instances. This can come about due to simple perceptual mistakes of the sort made by Kondabolu: he misheard a remark at the comedy club. But such errors are comparatively unusual, so it will help to list some other cases where the supposed over-identification of instances of unfairness or injustice instead stem from an agent's contested counterfactual judgments.

- A woman, Professor W, is talking to a man at a party. After introducing himself, he asks her what she does, and she explains that she's a professor of English literature, specializing in Shakespeare. Despite having no training in the field, he launches into a detailed explanation of the meaning of his favorite Shakespeare play, *King Lear*, muddling several key plot points. Later, recounting this incident to her husband, Professor W rolls her eyes at her conversational partner as a sexist mansplainer who most likely wouldn't have held forth in this way to a *male* English professor.[6] "You're being over-sensitive," her husband says: "He probably just

[5]A few methodological notes: while I concentrate on cases of racism and sexism or misogyny in this paper, for ease of exposition, similar cases featuring sensitivity to other forms of systemic oppression—e.g., classism, ableism, homophobia, transphobia, fatphobia, ageism, and so on—could easily be devised, and would, I believe, lend themselves to a similar analysis. Moreover, while many of my examples will—again for ease of exposition—be cases which predominantly involve *either* racism or sexism/misogyny, this is obviously compatible with recognizing that these are intersecting systems of oppression which are inseparable in both practice and (wider) theory. For two classic works on intersectionality, a concept developed by Kimberlé W. Crenshaw to do justice to the unique plight of Black women in the US, inter alia, see her (1991) and (1993) pieces. And for the general distinction between sexism and misogyny, and an intersectional approach to defining these concepts within the analytic feminist philosophical tradition, see my (2017), chap. 2, "Ameliorating misogyny" and chap. 3, "Discriminating sexism."

[6]See Rebecca Solnit, "Men explain things to me," reprinted in *Guernica* magazine, August 20, 2012 (https://www.guernicamag.com/rebecca-solnit-men-explain-things-to-me), which introduced the concept

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

jumped at the chance to talk about a play he loved." The implication being that he would have acted the same way regardless of his interlocutor's gender.

- Melody takes her daughter, Ellen, 4, out to the park, where Ellen has a meltdown over getting the wrong color Popsicle out of the cooler. "What a little spitfire!" remarks a passerby. "You've certainly got your hands full with that one." Melody is struck by the gendered overtones of that remark, since she doubts that a boy of the same age having a meltdown would have been labeled a "spitfire." Over dinner, her partner disagrees: he thinks that while subtly gendered terms like that certainly exist (e.g., "sassy"), she's being over-sensitive in this case: a boy would also have been called a spitfire for yelling "You're ruining my life!" loudly at a parent proffering the erroneously colored cold treat.
- Brianna, who is in a bad mood, cuts off her father-in-law abruptly during a long and boring story: "Where is this going?" she asks him, sighing audibly. He reacts to this slight with a fury that strikes her as being disproportionate and likely due to gendered norms of care—she can't imagine him getting that mad at his *son*-in-law, Bryn, for not listening attentively to his ramblings. When Brianna brings this up later to her husband, John, he thinks Brianna is being over-sensitive: his father has just had a short fuse with everyone lately.

These cases, drawn from real ones recounted to me, all involve counterfactual judgments on the part of a woman: to the effect that a man or boy would likely have been treated differently than either she herself or a female counterpart (due to a gendered norm of entitlement, a distinctively gendered term, and disproportionate punishment for women, respectively).[7] In all three cases, a man's disputing that counterfactual judgment furnishes him with the grounds, or at least evidence, for an accusation of over-sensitivity on her part.

I've chosen these three examples because the woman charged with over-sensitivity is not clearly right, and the man leveling this charge is not obviously wrong, about what actually went on in these cases. Still, his accusation of over-sensitivity toward her invites a number of more or less obvious responses that are worth distinguishing at the outset. To wit:

- The "I was there" response: in all of these cases, the woman actually witnessed the offending behavior (the putative mansplaining, the seemingly gendered remark, and the ostensibly outsized reaction due to a gendered norm violation). She may thereby have been privy to nuances of tone and inflection, and other contextual clues, that serve as evidence for her suspicions. Some of these will be difficult or impossible to articulate after the fact: our reading of a complex social situation may be accurate but inchoate. And even if we modify these scenarios slightly, such that the man accusing the woman of over-sensitivity was *also* present at the time, he may have missed these features, or been paying less attention to them. After all, subtle signs of sexism and misogyny constitute important social information that

of mansplaining (if not the term itself, which had a more diffuse origin on the internet); see also my (2020), chap. 8, "Unassuming: On the entitlement to knowledge."

[7]See my (2017), chaps 2–4. The fact that all of these three examples involve men behaving badly makes them easier to imagine but is arguably inessential—women can perpetuate misogyny and sexism too, fairly obviously, as my approach to the subject freely acknowledges and attempts to illuminate.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

she will typically be more invested in—and hence accustomed to—taking in, given its importance for surviving, let alone thriving, in a patriarchal social world.

- The evidence for the relevant counterfactual: In some cases, the woman may be able to point to specific evidence to back up her contentious counterfactual judgment. For example, Professor W might point out that *many* men have mansplained Shakespeare to her in the past when given the opportunity; the same is not true, anecdotally, for her male colleagues. Or Brianna might point out that her father-in-law has a long track record of subtly sexist and misogynistic behavior in a similar vein. For example: expecting his female relatives to clear the table and not contributing himself; dominating conversations in which women have as much or more to contribute; praising his granddaughters and grandsons in different, and gendered, ways.
- Evidence of personal reliability: In (again) some such cases, the woman may be able to point to the fact that she is typically reliable in picking up on sexism and misogyny when it is indeed present, yet doesn't go around making these sorts of accusations willy-nilly, or baselessly, whenever there is the relevant kind of social friction. (After all, she would certainly be punished heavily for so doing.) As Kondabolu remarked in the case with which I opened: "I have never been wrong before about [racism]!" While slightly—and deliberately—hyperbolic, weaker claims in this vein are common and apt when facing charges of over-identifying instances of sexism, misogyny, racism, and so on.

All of these are potential lines of defense to the effect that the woman accused of over-sensitivity is likely to be right and the man accusing her is likely to be wrong, in a given case. They may be more or less applicable; and they may or may not be accurate (again, I am leaving it open that the woman *is* actually wrong in any of these cases). But there's another point about the accusation of over-sensitivity that needs to be made here too. "Over-sensitivity" is a general accusation, typically implying the existence of a pattern: namely, that the woman is not just wrong in this case, but is too prone to "false alarms" in identifying instances of sexism or misogyny in general.[8] Given this, we might wonder: what about the risk of *under-sensitivity* in the sense of missing real instances of the relevant phenomenon? Why, given this omnipresent risk, does the prospect of over-sensitivity tend to dominate the discourse?

The comparison with other kinds of inquiry in messy, real-world contexts is instructive. When it comes to diagnostic testing, there is typically a trade-off between tests that have a higher percentage of false positives—or false alarms—and tests that have a higher percentage of false negatives—or misses. (In statistical terms, the former metric is sensitivity; the latter, specificity.) And in many endeavors, we are willing to tolerate a somewhat heightened risk of false alarms in order to avoid missing important phenomena. A rapid test for COVID that produces a higher percentage of false positives than a rival test might still be valuable if it misses a lower percentage of cases. An algorithm for detecting credit card fraud that "pings" with false alarms more frequently might still be a good one if it is also better at not letting real instances of fraud go under the radar. A sampling method for detecting spoiled food might well be useful even if it results in some perfectly good food being flagged in a higher percentage of cases, if it is less liable to miss potential food-borne pathogens.[9] The value of a detection mechanism

[8]Consider my construal of sensitivity, over-sensitivity, and under-sensitivity as *dispositions* in the introduction.

[9]My thanks to Daniel Greco and Derek Baker for discussion here.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

will depend on a complex set of considerations: not only these trade-offs, but the relevant costs and benefits thereof, together with information about base rates and prevalence. The point is just this: a higher proportion of false alarms is not inevitably alarming. Why aren't we similarly willing to tolerate the risk of false positives in order to detect more sexism, misogyny, and racism?

The obvious answer has to do with what we *do* with the relevant suspicions, and the prospect—salient to many minds—of false accusations in the latter case. Flagging food as spoiled when it isn't is one thing; flagging a *person* as a bigot, quite another. But we need to back up here: in the three cases above, and many others of this kind, there is not necessarily any prospect of accusations—false or otherwise—in the offing. That is, there may be no thought of confronting the partygoer, the passerby, or the father-in-law: in part because the threshold of certainty has not been met for such action. Perhaps the woman simply wants to avoid the offending party going forward.

We should not forget that, in some cases, the stakes for *not* avoiding people suspected of sexism and misogyny are high. Suppose that, in addition to seemingly mansplaining Shakespeare to her, Professor W's conversation partner at the party strikes her as being a little creepy: touching her arm and making excessively familiar comments, etc. Professor W hence decides to extricate herself quickly from the conversation and politely declines a ride home with him. If she was mistaken about the vibes, he has not been accused or wronged in any serious way;[10] if she was right to worry, she may have avoided calamity. (The spoiled food analogy is distressingly apt here.)

In many other cases, a woman will take no action whatsoever with respect to a party she suspects of sexism or misogyny. In each of the three preceding examples, she may simply want to vent her suspicions and frustrations, and note the appearance of gendered patterns that dog her and other women or girls throughout our lives, to a sympathetic audience. Who, as it turns out, is not only not sympathetic but levels an accusation in *her* direction: that she is being over-sensitive. It's telling that the risk of *this* accusation being false isn't as salient to many.

There is also the crucial point that, even when such accusations *are* leveled, they need not be of bigotry: they may be for the subtler and more common failing of channeling or perpetuating, say, sexist mores. It's not surprising if the father-in-law, given his gender and generation, has internalized the sense that women in particular owe him a sympathetic audience. Pointing this out could be done gently and, with a conscientious person, well-taken—or at least taken under advisement.

Finally, accusations of outright bigotry will, in some cases, be justified, even on the basis of imperfect information. It's not, after all, as if interpersonal interactions take place in a court of law: innocent until proven guilty is a legal standard, not a moral one. Imagine a case where a pattern of similar incidents unfolds with Brianna's father-in-law, and her evidence for his sexism becomes significantly stronger (though still admittedly imperfect). Calling him out for being sexist is not a legal proceeding, and he is not about to be arrested (or "canceled") on this basis. Still less is he being summarily thrown in prison. As in legal proceedings, a critical discussion can and should take place: is the accusation fair? Is the evidence sufficiently compelling? And, if it is, what if anything should be done on this basis? Given the preponderance of defenders and defenses a seemingly sexist or misogynistic agent will tend to have within a patriarchal society, it may again be valuable to have some others more willing to issue call-outs. We sometimes

[10]Cf. Basu (2019) on what we owe to each other, epistemically.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

have importantly different epistemic and moral *roles* within our community: a socially sensitive agent may play one such.

Consider, in this vein, Olivia Bailey's nice case of an agent, Dolores, who is exquisitely sensitive to the plight of the underdog, and the hurt of the outcast, in society.

> She is immediately and powerfully emotionally responsive to the moral ugliness of ostracism … Her acute sensitivity to this kind of hurt is sometimes admired by others: she manages to spot subtle forms of exclusion that they would never have caught if left to their own devices, but that they come to acknowledge as grounds for significant grievance with her help. (Bailey 2023).

At the same time, Dolores can be too quick to jump to conclusions: some of what she perceives as exclusion or ostracism is really just freedom of association. As Bailey points out, if Dolores undergoes a shift toward greater overall accuracy in her moral sensibility, something will have been gained, but something important will also have been lost. There is hence a case for ambivalence about this kind of moral improvement, Bailey argues, the overall assessment of which

> must depend upon the ways in which the moral community has come to depend on Dolores's underdog-oriented sensibility. If Dolores is known and counted on for her searing (if imperfectly nuanced) sensitivity … that is rare in her world, then moving away from her admittedly imperfect prior sensibility will amount to letting others down, at least for a time. The situation would be rather different if no one looked to Dolores in particular for that sensitivity. (Bailey 2023).

I agree but would add the perhaps more radical point that the question of whether to count the shift toward accuracy—which Bailey glosses as greater attunement to another real and competing moral value—*as* moral improvement often seems to me to be open. In statistical terms: even if Dolores's loss of sensitivity is more than made up for by gains in specificity, there are questions of base rates and the stakes of various kinds of error to consider. She may now be missing a great deal, and what she is missing may matter greatly, in an environment in which ostracism and exclusion are both common and deleterious. Moreover, as Bailey's discussion brings out, human beings are of course *not* just diagnostic tests or signal-detectors: part of what we are assessing, when it comes to the epistemic and moral value of someone's particular sensibility, is how much they can contribute to critical discussions following the alarm bells that sensitive agents are particularly adept at raising.

Moreover, such sensitivity can be a crucial survival skill—and a coping mechanism—for people facing oppression in virtue of their gender, race, and so on. It can help us advocate on our *own* behalf in ways that are often undermined by dominantly situated agents. In another incident Hari Kondabolu recounts, a group of three white men attacked him during his college years in Maine, yelling "What are you doing here? Why are you here?" repeatedly. Given Kondabolu's racially marginalized status as a brown man, and the white-dominated social environment, he clocked it as the racist attack it surely was. But instead of affirming his impressions, and reflecting on the relevant counterfactual—how much less likely he would have been to be attacked in this way as another white man in that environment—the white friends whom he sought sympathy from in the aftermath raced to question his impressions. According to one of them: "You

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

don't actually know it was a hate crime." "Did they use slurs?" "I don't think it's fair you're accusing them of a hate crime." Kondabolu: "I was shocked, because I didn't know he was their defense attorney." His friend's reaction is a nice, negative encapsulation of many of the points I've been making in this section.

There's also the background point that, such extreme incidents aside, racism is *not* just a collection of sins committed by bigoted individuals. Critical race theorists, following Charles Lawrence III, have long argued that racism is a kind of social disease faced by people of color and perpetuated by all of us—white people especially, given our investments in ignorance, denialism, and upholding our own privilege. (Kondabolu's friends in the above example are a clear case in point here; and, tellingly, he said their disbelief was as terrible as the attack itself.) Racism is hence well-understood on what Lawrence calls the epidemiological model, and addressed as an urgent public health matter (1987). Sexism and misogyny can be productively understood in a similar vein, according to my own adaptation of Lawrence's insights (see, e.g., Manne 2017, 28). A corollary which has emerged in this section: the presence of some individuals who may be liable to a greater percentage of false positives than their less sensitive counterparts is potentially highly valuable in detecting instances of sexism, misogyny, and racism when it is occurring. To halt the progression of these social diseases, we need people who miss very little, even if they are occasionally mistaken.

Are these people sensitive, or slightly over-sensitive albeit in a defensible and valuable way? The answer to this question will depend on how short the agent falls of the threshold for "sufficient reliability" that my definition for sensitivity has built into it. We can thus imagine two ways of reading the Kondabolu case with which I opened: first, he is disposed to correctly identify racism with sufficient reliability to count as sensitive, not over-sensitive (or "too sensitive"), per his own diagnosis. He just unluckily got it wrong on this particular occasion.[11] Second, and more subtly, he may indeed be slightly over-sensitive to the relevant phenomenon—disposed to identify instances of racism slightly more often than would qualify him as "just sensitive"—but in a way that is a valuable corrective to the under-sensitivity of other agents in his epistemic and moral community.

2. Over-extension of concepts

We can now move to consider the second kind of supposed over-sensitivity I want to discuss in this paper: conceptual expansion, where counting novel phenomena under headings like sexism, misogyny, and racism garner the accusation that people are becoming over-sensitive. The psychologist Nick Haslam articulates this worry in a well-known recent paper, "Concept creep: Psychology's expanding concepts." Haslam argues that certain, negative concepts—including abuse, trauma, bullying, and prejudice—have been expanding in both "vertical" and "horizontal" directions over the past few decades. Roughly, vertical expansion involves lowering the threshold for counting phenomena as falling under the extension of a concept—counting comparatively mild teasing or physical "joshing" as bullying, for example, or loosening the requirement that bullying

[11]On other occasions, the mistake may be due to other factors: as a brown man, he might attribute a store clerk's unpleasant attitude to racism when they are in fact unpleasant to all of their customers, for instance. Notice that the attribution of sensitivity to an agent allows that he may get it wrong occasionally, as long as his disposition meets the threshold for reliability on my definition, even when the source of the mistake is somewhat more patterned or predictable than a simple perceptual error.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

behavior has to be repeated (Haslam 2016, 4–6). Or it may, claims Haslam, be attributable to shifting the criteria for the application of a concept from the perspective of the perpetrator to the perspective of the victim, thereby supposedly introducing an element of subjectivity into its extension. Counting emotional abuse as such requires taking into account how some form of treatment *felt* to the recipient, and not just what was done or how it was intended (Haslam 2016, 3, 11). Horizontal expansion involves enlarging the extension of a concept to include adjacent phenomena—recognizing cyber-bullying, for example. It may also, adds Haslam, count inactions or omissions in with actions or commissions—counting social exclusion as bullying, for instance, in addition to acts of violence, threats, intimidation, and spreading rumors about the target (Haslam 2016, 4–6). "Conceptual creep has occurred across a diverse assortment of concepts and has commonly involved an increased sensitivity to negative experience and behavior, an increased focus on harmful forms of inaction, and an increased acceptance of subjective criteria for deciding when the concepts apply," Haslam summarizes (2016, 10).

Examining Haslam's case study of concept creep with respect to prejudice, and his worries about the drivers and results of this, will be particularly instructive given my aims in this paper. He traces the study of this topic to Gordon Allport's 1954 *The nature of prejudice*, which defined prejudice as "an antipathy based upon a faulty and inflexible generalization. It may be directed toward a group as a whole, or toward an individual because he [sic] is a member of that group." (Allport 1954, 9). Haslam notes that subsequent scholars were not content to leave things there, introducing concepts like aversive prejudice, where instead of harboring conscious antipathy toward a group, people are instead fearful, uncomfortable, or uneasy around its members, sometimes without any conscious awareness of being so. The related concept of implicit bias incorporates the idea of harboring unfair negative associations with a certain group, among other things, again without conscious awareness. Concepts like "benevolent sexism" crystallize the idea that some kinds of prejudice involve superficially good, but patronizing, understandings of certain groups of people: stereotyping women as caring, innocent, and lacking in aggression, say. Finally, concepts like micro-aggressions count subtler and less obvious slights toward a social group as instances of prejudice (Haslam 2016, 9–10).

Haslam's notion of concept creep is ostensibly descriptive, not normative. "It is possible to define 'prejudice' in a coherent way that encompasses all of these variants," he acknowledges. "My point is simply that the concept now refers to much more than it did several decades ago." (Haslam 2016, 10). Moreover: "Each extension of the concept of prejudice was arguably well justified." (Haslam 2016, 10). And yet the slippage to a critical mode is soon easily discernible, arguably already present in the tendentious label "creep" (as opposed to the more neutral label "expansion") suggestive as it is of an undeliberate, insidious, ooze.

Haslam considers various explanations for concept creep, and takes the most promising to be Steven Pinker's idea that minority groups have created various "rights revolutions," whose campaigns "are propelled by an escalating sensitivity to new forms of harm"—some of which he is explicitly critical of as over-sensitivity—and "the slightest trace of a mindset that might lead to [them]" (Pinker 2011, cited in Haslam 2016, 13). Haslam: "A general sensitization to harm could go a long way toward accounting for concept creep." (2016, 13). More: "Psychology has played a role in the liberal agenda of sensitivity to harm and responsiveness to the harmed." (Haslam 2016, 13). Evincing what he describes as a more ambivalent attitude toward the resulting concept creep, he worries

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

that it has created a new culture of victimhood, in which people are typecast as various kinds of victims on the one hand, and moral villains on the other: abusers, traumatizers, bullies, and bigots. As the result of this, as well as the lower threshold for applying the relevant concepts ("semantic dilution"), "concept creep may release a flood of unjustified accusations and litigation, as well as excessive and disproportionate enforcement regimes" (Haslam 2016, 14). Moreover, "concept creep runs the risk of creating a public impression of psychology as a field that exaggerates misery . . . and is over-sensitive to perceived bias and discrimination."[12] The upshot of Haslam's discussion is this: concept creep with respect to prejudice stems in large part from over-sensitivity. And it is liable to result in false accusations.

Haslam's picture is in several respects seriously misleading. Allport was hardly the first theorist of prejudice in general and racism in particular. Even confining our focus to the Anglo-American context, Black scholars like Frederick Douglass had been theorizing about these matters for well over a century (1845/1979).[13] Some fifty years later, in 1898, W. E. B. Du Bois analyzed racial prejudice as the conviction "that people of Negro blood should not be admitted into the group life of the nation no matter what their condition may be" (Du Bois 2015, 82), already pointing to some of the structural sources of inequality that Haslam does not reference, including economic disadvantage. In Du Bois's famous notion of double-consciousness, one will have as a Black subject a "sense of always looking at oneself through the eyes of others, of measuring one's soul by the tape of a world that looks on in amused contempt and pity" (2007, 3). Here we have clear precedent for the supposedly recent ideas of prejudice not always involving antipathy, and instead being patronizing or pitying.

Meanwhile, feminist theorists have been clear about the fact that sexist prejudice may involve superficially flattering stereotypes about women, which actually serve to control and disempower us, since at least Mary Wollstonecraft. In her *A vindication of the rights of women* (1792), she wrote in opening:

> My own sex, I hope, will excuse me, if I treat them like rational creatures, instead of flattering their FASCINATING graces, and viewing them as if they were in a state of perpetual childhood, unable to stand alone. I earnestly wish to point out in what true dignity and human happiness consists—I wish to persuade women to endeavour to acquire strength, both of mind and body, and to convince them, that the soft phrases, susceptibility of heart, delicacy of sentiment, and refinement of taste, are almost synonymous with epithets of weakness, and that those beings who are only the objects of pity and that kind of love, which has been termed its sister, will soon become objects of contempt.
>
> Dismissing then those pretty feminine phrases, which the men condescendingly use to soften our slavish dependence, and despising that weak elegancy of mind, exquisite sensibility, and sweet docility of manners, supposed to be the sexual characteristics of the weaker vessel, I wish to show that elegance is inferior to virtue,

[12]Haslam 2016, 15. As this paragraph brings out, Haslam's article toggles between considering concept creep as a broad social and cultural phenomenon and as one that primarily afflicts and has happened within psychology. Given my purposes in this paper, I focus on the former, broader concern here, which I take to be the more natural one—it would be strange to worry about concept creep in psychology in particular or in the academy in general but not at all beyond it.

[13]For a sophisticated treatment of even older precedents, see Jorati (2023).

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

> that the first object of laudable ambition is to obtain a character as a human being, regardless of the distinction of sex. (1792, Introduction)

Wollstonecraft is not alone here: it is hardly an exaggeration to say that noticing the insidious and patronizing quality of superficially positive stereotypes about women animated much of early feminist theorizing.[14]

To state the obvious, then: this is not a matter of there once having been a narrower concept of prejudice which was enlarged, rightly or wrongly, in recent decades, due to heightened sensitivity to harm. Rather, a white man, widely treated as an authority on the topic, missed a great deal that people directly affected by prejudice—Black people and white women among them—had long been theorizing, with limited uptake in the academy. Their ideas were then gradually rediscovered, or reinvented, by subsequent scholars.

That subordinate group members would typically be the people to break ground on concepts of prejudice—which affect them the most directly and acutely—is deeply unsurprising. My point here does not rely on any particularly controversial claims within standpoint epistemology. As Lidal Dror has recently argued, there is widespread agreement with the idea that socially marginalized people, in virtue of their social position, "*tend* to have a superior epistemic position than non-oppressed people when it comes to knowing things about the workings of social marginalization that concern them." This is because they "*tend* to have more relevant experience and motivation" to acquire such knowledge (his emphasis). Dror calls this the weak inversion thesis, and endorses it—while ultimately arguing against the stronger inversion thesis that oppressed people have this advantage *in principle* (2023, 624). I do not need such a strong claim for my purposes either. For all I have said, the epistemic advantage of people of color and marginalized genders may consist in contingent tendencies to be more sensitive to phenomena that affect us more in general, but are in principle open to people in any social position to acquire knowledge of by paying careful attention to the relevant evidence (including testimonial sources). To adapt Elizabeth Anderson's (2020) example: you don't need to be a standpoint epistemologist to acknowledge that car mechanics, who are tasked with figuring out problems with cars and fixing them for their livelihood, are likely to be the most insightful people when it comes to car repairs and maintenance. Similarly, I think it is hard to deny that the people for whom

[14]Similar ideas also found development in the work of Harriet Taylor Mill (sometimes writing with John Stuart Mill, her husband), in the nineteenth century. See, e.g., Taylor Mill's "Enfranchisement of Women," on the alleged hardening influence of women's enfranchisement, and the way a positive emphasis on women's "softness" is tantamount to keeping women weak in both body and mind and, hence, ultimately, subordinate to their husbands (2022, 26–27). Another obvious missing reference is Simone de Beauvoir's *The second sex* (published in 1949 as *Le deuxième sexe*). See, in particular, her remarks on "the eternal feminine," and the tropes of the saintly mother and the angelic young girl (which contrast with the cruel stepmother and the perverse virgin, respectively). Beauvoir writes, further, that "to identify Woman with Altruism is to guarantee man absolute rights to her devotion; it is to impose on women a categorical must-be." (2011, 267–68). However, as these remarks unwittingly serve to bring out, it is important to note that stereotypes about women's "softness" and similar often exclude Black women in particular, and so cannot be taken to be universal in the ways that some feminists have erroneously taken them to be. Consider, in this connection, Sojourner Truth's famous "Ain't I a woman?" speech, which attacks the idea that her strength, agency, and perceived lack of softness makes her ineligible for womanhood (see https://www.nps.gov/articles/sojourner-truth.html for the standard reconstruction of her extemporaneous remarks). I thank an anonymous referee for bringing this important matter to my attention.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

surviving, let alone thriving, depends on understanding the contours of prejudice will typically be the most insightful when it comes to first theorizing these concepts.

What Haslam terms "concept creep" with respect to prejudice is thus far not alarming, or even an apt source of ambivalence. It is also not obviously true that wider understandings of such concepts are due to greater sensitivity to harm per se, as opposed to the specific harms of racism, sexism, misogyny, and so on. This is not obviously a problem, even when it does stem from and result in increased sensitivity. Take the concept of micro-aggressions, for example, which are, roughly speaking, slights that have the potential to accumulate and add up to something seriously harmful for their targets or victims.[15] Plausibly, this concept *has* resulted from an increased sensitivity to the ills of structural oppression and how they can be manifested. And it has enabled oppressed people to name harms and wrongs they thereby face that would otherwise be very difficult to identify: how do you complain of a "death by a thousand cuts" without being empowered to point out the scratches?

This leaves us to consider the central concern driving those up in arms against "concept creep": the risk of false accusations and increased moral litigation of various harms and wrongs. But here again Haslam fails to convince: many of the ways of expanding the concept of prejudice that he himself highlights deliberately militate in the opposite direction.

Take the concept of implicit bias. The idea that many of us—those who face oppression, included—harbor implicit racial biases, over which we have no conscious awareness and subsequently little control, has long been taken to reduce the aptness of accusations of racism. "Implicit race bias does not mean a person is a racist," write the psychologists R. Banaji, Susan Fiske, and Douglas Massey (2021).[16] Even as a concept expands within the academy, we can make important distinctions: between someone's channeling or perpetuating racist social forces, and someone's deserving the pejorative label "racist" per se. Similarly, in my own treatment of misogyny, I argue that there can be misogyny without misogynists, and that those agents who *do* deserve the label "misogynist" have to exceed a certain threshold of severity and repetition of misogynistic offenses. "Misogynist" can act as a kind of warning label, then, to highlight the worst offenders within a particular social environment (Manne 2017, 65–66). The risk of "vertical inflation" of a kind that might give rise to false accusations is checked by design—and definition.

This leaves Haslam's worry that defining concepts of prejudice partly from the perspective of the target or victim, rather than the perspective of the perpetrator, introduces a troubling element of subjectivity. Again, I think this worry is at least overstated. Take, again, my concept of misogyny, which makes just such a shift, and holds that misogyny is hostility and hatred girls and women *face* rather than something boys and men feel, at least in the first instance. I subsequently take misogyny to be, for a rough first pass, the hostile and hateful treatment girls and women face that serves to police and enforce a patriarchal social order (Manne 2017, 33–34). I advocate this analysis of the concept for a number of key reasons: first, that the naïve conception of

[15]For state-of-the-art treatments of microaggressions within philosophy, see Weekes Schroer and Freeman (2019) and McClure and Rini (2020).

[16]Cf. Holroyd (2012) on implicit bias and moral responsibility—one of many important treatments of various aspects of the subject in analytic feminist philosophy. See, e.g., Brownstein and Saul (2016) and Beeghly and Madva (2020) for two state-of-the-art edited volumes on implicit bias, each of which offers excellent critical introductions by these authors.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

misogyny, according to which it comprises hostility and hatred toward all women, or even women generally, is rarely going to be instantiated in a patriarchal society in which women are socialized to please and serve men. Why would they bite (or hate) the hands that feed and serve them? (Manne 2017, 46–48). Moreover, construing misogyny as hatred in men's hearts makes it not only a psychological oddity but an epistemological quagmire: it is hard to know what someone feels, deep down, short of being their intimate relation or therapist (Manne 2017, 44). The naïve concept of misogyny thus deprives women of a morally powerful tool to name the problems facing us: better to reserve the term for a more important concept, I argue.

Crucially, however, what counts on my treatment as hostile and hateful treatment depends upon the "reasonable woman" standard developed by Catharine MacKinnon in the context of sexual harassment law (MacKinnon 1979). According to this conception, not just anything goes when it comes to how a certain form of treatment "comes across" to the target or victim thereof: she must be reacting in a reasonable way given the particular details of her social position—including, crucially, her gender (Manne 2017, 60–61). It's not just that women *feel* intimidated or attacked by seeing pornographic posters pinned up in their workplace, for example; it's that it is reasonable to feel this way as a woman for whom sexually objectifying imagery is fraught with social meaning.

The upshot of this section is that the charges of over-sensitivity with respect to expanding notions of sexism, misogyny, racism, and the like, as crystallized by the literature on "concept creep," tend to be unwarranted. Typically, it is less a matter of these concepts being injudiciously expanded as that an artificially narrow concept, delineated by a dominant group member who was wrongly treated as an authority figure on the subject, was recognized as such. Moreover, sometimes concepts *should* be expanded in response to an increased sensitivity to what oppressed people are actually up against. Finally, fears about moral litigation due to "concept creep" are at least overblown: expanding the concept of racism, sexism, and misogyny in the ways that have been advocated by anti-racist and feminist scholars has characteristically led to a *less* punitive and individualistic approach to the subject. The "concept cops," as we might call them, are sounding false alarm bells.

3. Overreactions in feelings

Now to consider the third kind of alleged over-sensitivity to sexism, misogyny, racism, and so on, in which people are accused not of misidentifying instances, or operating with too expansive a concept of the relevant harms, but are instead held to be reacting too strongly to them, emotionally. They are held, in other words, to be "snowflakes."

I want to emphasize that overreactions are a real thing, and I do not aim here to exonerate—or give *carte blanche* to—oppressed people accused of them. Still less do I want to do so with respect to people accused of overreacting in general. Take, for example, Brianna's father-in-law from my example in the first section, who overreacted angrily to a minor moment of conversational frisson. Or take, for a more politically charged case, those whites charged with "white fragility," to use Robin DiAngelo's (2018) term. These are white people who react to any attempt to have them face the realities of racism with denialism, incredulity, and, crucially for my purposes here, tears, rage, and so on. One doesn't have to be on board with every aspect of DiAngelo's approach—or every exercise promoted by her and her ilk in corporate anti-racism trainings—to recognize the basic truth of her point here: white people often react to discussions of race and racism with excessively big feelings. Hence, any sensible discussion of emotional

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

overreactions in general, and to socially charged subjects in particular, will have to be indexed to substantive normative judgments. There is no getting away from these whenever we have to make calls about which emotions are apt, fitting, warranted, and so on; what one is entitled or even obligated to feel depends, among other matters, on how good or bad things are and to what extent they ought to be or be otherwise. I proceed from the assumption—hopefully uncontroversial among my readers—that sexism, misogyny, and racism are very bad indeed and urgently need remedying.

So return to the moment with which I opened, of Kondabolu's anger at the remark he misheard at the comedy club. Specifically, he threatened the heckler: "Why don't you come up here so I can kick your teeth in?" Even if he *had* been right about this being an instance of racism, calling this an overreaction without further comment will seem natural to many people. But I think it is mistaken.

In her book-length defense of the value of rage, the philosopher Myisha Cherry points out that the insistence that anger or rage be proportionate to the wrong or offense suffered requires of oppressed people that they react perfectly—and keep reacting perfectly, with no more or less emotion than is apt—in every individual instance.[17] What is far more common is what happened in Kondabolu's case: he took racist heckling on the chin umpteen times, and simply ignored it onstage for years, until that one fateful evening. The heckling on that occasion was the proverbial straw that broke the camel's back. (He was subject to moral bad luck, in that *this* was the time he misheard the utterance.) I do not think demanding unbreakable fortitude is morally reasonable.

Again, we can distinguish here between two possibilities: one is that someone like Kondabolu is *not* overreacting: had someone actually said "Palestinian power!" in order to mock him, he would have been justified in feeling fury (especially after keeping his temper about other, similar incidents).[18] Alternatively, we might say that Kondabolu *was* overreacting, but that this overreaction is understandable and not blameworthy in light of all he had been dealing with, for many years, in many contexts.[19]

I am tempted to think that both responses have it right, but regarding different intentional objects of the agent's anger.[20] Considered as a response to this particular racist incident, it *is* an overreaction, albeit an understandable and not blameworthy one given the social context. But considered as a response to a whole string of similar incidents, it is *not* an overreaction: it is in fact warranted, even if it is also partly misdirected (or, rather, concentrated on just one of the many people who the agent is ultimately angry with).[21] Lest I be misunderstood here, recall the object of our appraisal: the *feeling*, not the threatening remark (still less any action undertaken on that basis).

[17]Cherry (2021, 48–50); cf. Srinivasan (2018, 130), who argues that, in order to be apt, a subject's anger must "be properly motivated by that possessed reason [for anger], and proportional to that reason … [S]uppose I find out that you lied to me about liking a meal I cooked, and I am thereafter engulfed by a wild and lifelong fury." This anger is not apt, claims Srinivasan, because the "anger is disproportionate to [the] reason." See also Tessman (2005), chap. 5, "The burden of political resistance," for nuanced discussion here.

[18]Remember that my definition of sensitivity builds in that the reaction should be fitting given the *putative* truth of the relevant judgment: so the perceptual error in question here can be bracketed for the purposes of the discussion.

[19]Compare Tessman (2005) on a related form of moral sensitivity, namely, to the plight of suffering others, which can lead to unsustainable emotional anguish; chap. 4, "Between indifference and anguish."

[20]Thanks to Zev Friedlaender and Justin D'Arms for valuable discussion here.

[21]Compare Cherry: "it's hard to judge if having a strong response to a microaggression would actually be an overreaction. It could be one microaggression after an accumulation of others, and the accumulation could give someone good reason to feel rage that is large in its level and intensity." (2021, 49).

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

That is, I certainly don't believe that Kondabolu would have been justified in actually kicking someone's teeth in, or even that he was justified in yelling at them threateningly. (This is a verdict that he himself evidently agrees with: he called his remark the worst and harshest thing he'd ever said to another human being, obviously regretful.) But demonizing his feelings seems to me wrongheaded.

As we saw in section 1, another reason to be careful of demonizing an agent's over-sensitivity to systemic injustice is that under-sensitivity to these features of the world is a large and common problem. This holds with respect to overreactions to systemic injustice too, partly for the obvious reason that many non-oppressed people are prone to affective under-reactions to misogyny, racism, and so on. But it holds for a more surprising and interesting reason too: such under-reactions are *also* a predictable outcome of certain forms of social subordination faced by the people subject to oppression.

To see why, consider Kate Hamilton's harrowing memoir, *Mad wife*, in which she recounts her experiences of being sexually violated for many years by her husband. She describes acquiescing night after night without true consent, without desire, certainly without pleasure. Her devil's bargain: "You can fuck me, but you can't kiss me and I don't have to pretend to like it." (Hamilton 2024, 8). Here is how she describes a typical evening:

> I'm reading in bed. This night, I have not heard his footsteps on the stairs in time to turn the light off and feign sleep before his arrival in the bedroom. He undresses and gets into bed next to me, pressing his body alongside mine. I continue reading. He begins to caress my body. I do not look at him or acknowledge him in any way. His caresses become targeted . . . I do not want to have sex with him now and have not wanted to for a long time. He knows this. I ignore what he is doing as long as possible, focusing all my attention on my book . . . Now I've put the book down, one hand holding it across my chest in a kind of defense posture. My other arm lies still by my side. My head is turned away, lest he forget the no-kissing rule. I do not release the book and put my arms around him. I do not caress him. But I am not storming out: he continues. I need this to end as quickly as possible so I curl up in my mind, eyes shut, far away from the machine of my body. A fly on the ceiling would see a splayed, naked, immobile woman, head turned, eyes squeezed shut, face intensely focused, and a man doing urgent things to her body. (Hamilton 2024, 8–9).

Hamilton dragged her husband, despite his extreme reluctance, to therapy. There, mustering all her courage, she confesses what she describes as her deepest, darkest truth: "When we have sex, I feel like I'm being violated." He shrugged. "She's always so melodramatic." Hamilton wasn't surprised by this response. "It was his standard reaction to my complaints about the sad state of our marriage, his way of training me to see my needs—emotional connection and communication—as excessive, and his (primarily sex) as entirely reasonable." (Hamilton 2018).

Several lessons emerge from this case and others like it. First, fairly obviously, accusations of overreacting are frequently weaponized by dominant group members against more vulnerable people, including those over whom they have a great deal of power. This feeds into gaslighting: the insidious process by means of which the target is made to feel defective for harboring beliefs, desires, needs, feelings, and other mental states to which she is entitled[22]—here, her profound aversion toward this sex she did not

[22]Manne (2023, 23); see also Abramson (2014) for a classic and illuminating account of the moral harms of gaslighting.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

want, and her need for the bodily autonomy it took her years to claw her way back to. And although overreactions are of course a live possibility in any complex human domain in which emotions run high and denials go deep, under-reactions are very much a problem too, and may be at least as likely.

This isn't just a point about other agents who undermine our warranted reactions, and gaslight us in the process. We gaslight ourselves too, I think, in coming to doubt and mistrust our own initial reactions. The popular subreddit "amIoverreacting" contains no shortage of examples of this, where the fact that a woman is *not* overreacting to a man's bad behavior can be gleaned just from reading the following titles: "Am I overreacting that my partner keeps getting so drunk he wets the bed?"[23] "My brother said I shouldn't be mean to the man that abused us for years. Am I overreacting for being upset?"[24] "Am I overreacting [that my] boyfriend keeps putting his fingers inside me as a 'joke'?"[25] No, no, and no. The fact that these posters went to reddit to ask their questions tends to suggest that they *know* at some level that they are not overreacting. They just needed to be emboldened by other posters to feel what they were feeling, and perhaps to take appropriate action on this basis. Sometimes we need other people to tell us that we are in the right, not only morally but emotionally; another's anger on our behalf can jolt us out of self-doubt and vacillation.

There is also a lesson to be drawn from such cases about the precarity of sensitivity, and the ways people who are repeatedly wronged in serious ways—assaulted, threatened, violated, and so on—may become inured to their own predicament. There are already hints of the numbness Hamilton came to feel in her account of the sex itself: the dissociation in the moment, the "vacating" of her body, the becoming a "machine." And that coerced disembodiment—a common coping mechanism for victims of sexual abuse and assault—began to suffuse her life in general. "Submitting to unwanted sex with someone who knew it was unwanted made me feel—during the sex and, over time, in general—like a thing, a vessel for my husband's needs. A sex doll, made wholly of orifices. It made me feel like I was not a person. I no longer existed, not as a complete human being with needs, feelings, opinions, beliefs, agency." (Hamilton 2024, 10). Or, more simply: "I felt like a sex doll. I felt unselfed." (Hamilton 2018).

Hamilton describes the process of finding her way back to herself partly by getting in touch with a feeling she had long submerged: rage. "Way deep down, where I could not even see or feel it until nearly a decade later, in tandem with my revulsion grew my rage." (Hamilton 2024, 9).

Philosophy, with its emphasis on reason and rationality, has historically tended to err in a particular direction, overfocusing on constraining rage and policing other intense feelings. Many feminist theorists have recognized this point, and it has spurred a lively conversation—both inside and outside the academy—in recent years about the way anger and rage may be crucial tools for political mobilization, as well as simply apt responses to injustice and oppression.[26] But what is perhaps less recognized, and

[23]https://www.reddit.com/r/AmIOverreacting/comments/1ds5ugz/aio_that_my_partner_keeps_getting_so_drunk_he/.

[24]https://www.reddit.com/r/AmIOverreacting/comments/1dziw84/my_brother_said_i_shouldnt_be_mean_to_the_man/.

[25]https://www.reddit.com/r/AmIOverreacting/comments/1dxpzts/aio_bf_keeps_putting_his_fingers_inside_me_as_a/.

[26]See, e.g., Cooper (2018), Chemaly (2018), Traister (2018), Srinivasan (2018), and Cherry (2021). For a much more skeptical treatment of the value of prolonged anger, see Nussbaum (2016).

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

deserves our attention too, is that these emotions can be a precious resource that are hard to preserve. They can become difficult to cultivate in the first place for some people.

In her recent book, *To those who have confused you to be a person*, the journalist Alia Dastagir recounts the stories of women threatened online, and subject to misogynistic abuse and doxing, often due to their writing and advocacy on issues of social justice. Many of these women have become inured to it, and testify to dissociation in the moment, followed by emotional numbness in the aftermath. Taylor Lorenz, for example, was subject to numerous vicious mobbings for her reporting about the tech world and right-wing sources on social media. "It's like I cannot feel emotion. I've gone through so much that . . . I feel numb," she told Dastagir in an interview (Dastagir 2025, 116). The activist Jaclyn Friedman, a sexual assault survivor who wrote about her own experiences, spoke of the feeling of leaving her body during the online backlash. Disinformation researcher Abbie Richards recalled blacking out during the most intense attacks on her. When Dastagir interviewed me, I told her I wouldn't be able to help much: as she writes in the section of the book, "Dissociation, desensitization, numbing," I have almost no memory of the worst times of being subject to misogynistic abuse and vitriol (Dastagir 2025, 116).

Psychologists vary in their treatment of dissociation; it is sometimes theorized as the thing our brains do to protect us at the worst of times, sometimes as a very ordinary mechanism that explains the feeling of blankness any time we go into "autopilot mode" and lose track of our own movements.[27] (The sense of having driven a familiar route without really being aware of doing so is a common example of ordinary, benign dissociation in the literature.[28]) Whatever the case, it's worth noting that, when a woman experienced sexual assault, her dissociative symptoms one month later were predictive of subsequent poor psychological outcomes; so, interestingly, was her anger (Feeny et al. 2000).[29]

As Dastagir points out, it's impossible to know, without clinical evaluation as well as conceptual clarity, the extent to which dissociation was at stake in these cases and others like them, of misogynistic attacks on women who dared to be outspoken. It may be emotional numbing and cognitive avoidance too. And, as Dastagir argues, all of these responses can be useful coping mechanisms; they can also lead to harms, both short-term and long-term. In the short term, emotional numbness can compromise one's ability to detect serious threats and dangers. Dastagir interviewed Leah Torres, an OB-GYN who provided abortions in red states, and maintained a Twitter presence to explain her work and advocacy. Torres described the experience of having been brutally maligned and attacked over the years for tweeting things like "Do dead women from forced pregnancy matter?" (This was in 2013, in response to a tweet about dead babies.)

In 2018, someone on Twitter asked Torres if, when she laid down to sleep at night, she could hear the dead babies screaming. She finally snapped: "You know fetuses can't scream, right? I transect the cord 1st so there's really no opportunity, if they're even far

[27]For three standard discussions of the subject, see Steinberg and Schnall (2001), Simeon and Abugel (2008), and Dell and O'Neill (2009).

[28]See, e.g., David Spiegel, MD, "Overview of dissociative disorders," *Merck Manuals*, Reviewed/Revised May 2023, https://www.merckmanuals.com/home/mental-health-disorders/dissociative-disorders/overview-of-dissociative-disorders.

[29]Of course, these correlations should not be taken to imply causation, since both anger and dissociation may not cause but be the natural concomitants of difficulty coping with one of the most serious violations imaginable, for which society offers little supportive infrastructure and, on the contrary, often heaps blame upon the victim.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

enough along to have a larynx." The response was vitriolic, and ultimately cost her a job in Utah and a temporary medical license in Alabama, where she had moved to practice.[30] Torres: "You get numb. And then you can't assess threats correctly . . . You have to keep going. In order to combat violence, you have to fight your own experiences."[31] In the long term, such fighting of experience can compromise your ability to experience much of anything. Dastagir: "Emotional numbness is characterized by diminished interest, disconnection, and a lack of emotional response. It can feel like emptiness or a daze." When you are emotionally numb, you don't feel the worst things, "but you may not feel good things either. You may not care about things that once mattered to you. You may stop caring about yourself." (Dastagir 2025, 117). And such care, is of course, needed for surviving—again, let alone thriving and continuing to analyze, advocate, create. The poet Joy Sullivan recently put it this way: "I'm no stranger to the occasional internet hate, especially as my platform has grown and I've had poems take off on Instagram. Over the years, I've had to develop a thick skin. I realize fortitude is vital and that, in many ways, I need to toughen up. Yet, that kind of neutrality is tricky to balance against my primary task as a poet—to remain open and bruisable to this world."[32]

So to those who worry about emotional overreactions, I would counter: a disposition to respond with dismay to systemic oppression is in fact a precious resource that merits conservation. When this disposition meets the threshold for sufficient reliability, this is just sensitivity, per my definition, which need not be perfect in order to be valuable. But even when it falls somewhat short of this mark, and makes an agent more prone to emotional overreactions than we would like (i.e., she *is* over-sensitive in this sense), it is also a disposition that protects her from a real risk: becoming an *under-reactor* due to gaslighting or numbness. And given that we typically need to *feel* the badness of racism and misogyny in order to fight them effectively, this risk should be guarded against with care. Recalling the opening of this section, my aim is hence not to give *carte blanche* to people who overreact, or to condemn those who criticize overreactions as necessarily mistaken; a great deal will depend upon the context and the stakes, as well as the frequency and the nature of the overreaction in question. My plea is a considerably more modest one: to simply bear all of this in mind and not respond to overreactions too vehemently or hastily. Ironically, this would be an overreaction.

4. Conclusion

In this paper, I have argued that the supposed over-sensitivity of people nowadays with respect to racism, misogyny, and the like has been consistently exaggerated in both academic and popular discourse. There are three main things that could be meant by such over-sensitivity: first, that people are too likely to identify instances of racism, misogyny, and similar, due to their alarm bells sounding regarding innocent phenomena; second, that our concepts of prejudice are expanding in unhelpful directions; and third, that people's emotional reactions to putative instances of these phenomena

[30]Following investigation in part for making "public statements related to the practice of medicine which violate the high standards of honesty, diligence, prudence, and ethical integrity demanded from physicians licensed to practice in Alabama," Torres was eventually cleared and granted her permanent medical license; Dastagir 2025 59.

[31]Dr Leah Torres, personal communication.

[32]Joy Sullivan, "Woman behind a paywall," *Necessary Salt* Substack, July 12, 2024, https://joysullivan.substack.com/p/woman-behind-a-paywall.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

are problematically hyperbolic. In each of these cases, I have argued there are important reasons to be cautious about over-sensitivity diagnoses. In the first, an agent's supposed over-sensitivity to racism and misogyny is often simply sensitivity to phenomena which are routinely minimized; in other cases, her slight over-sensitivity nevertheless has value in a social environment in which under-sensitivity is a large and predictable problem. In the second case, the supposed over-expansion of the relevant concepts often corrects for their artificial narrowing, and additionally has an important role in making us more sensitive to genuine harms and wrongs. In the third case, an agent's supposed over-reactions to racism and misogyny may be warranted, and also reflect an affective disposition that constitutes a valuable defense against the twin perils of gaslighting and numbness.

In closing, though, it's worth considering whether the shoe may be on the other foot with respect to over-sensitivity: could dominant group members be seriously over-sensitive in their own right? Could they be overreacting to supposed harms and risks to their group? Could they be the true snowflakes?

I cannot hope to fully address this question in this paper, so what I say will, by necessity, be merely suggestive and programmatic. But it behooves us to notice that we find plausible examples of this ready to hand in the very discussion of the supposed over-sensitivity of oppressed people, in ways that mirror the three-part distinction that has formed the backbone of this paper. To start with: Nick Haslam highlights the risk of false accusations with respect to concepts like implicit bias and misogyny, even though these in fact discourage individualistic and punitive reactions. This seems on the face of it to be over-identification: registering "false positives" with respect to which accusations are being leveled, thus constituting the first kind of over-sensitivity I distinguished in this paper. Secondly, right-leaning individuals have, notoriously, expanded the concept of "racism" to encompass "reverse racism," wherein white people are held to be suffering from much the same kind of systemic prejudice that Black people have historically faced in American culture. This is, I take it, an unhelpful over-expansion of the relevant concept, and thus constitutes the second form of over-sensitivity discussed here. Finally, for an example of overreaction: consider Jonathan Haidt and Greg Lukianoff's "The coddling of the American mind" (2015a), the September 2015 cover story for *The Atlantic*, which argues that college students are becoming over-sensitive to potentially fraught material, particularly that which might trigger PTSD symptoms for previously traumatized individuals (military veterans, sexual assault victims, and so on). Their backstory was included in the sidebar of the magazine article: Haidt was worried that student complaints about his teaching might escalate and cost him his career.[33] Lukianoff had sunk into a deep depression because he felt at risk of being "canceled" for causing offense professionally. "I could not shake the feeling that any mistake I made in my professional life . . . would mean ruin; nightmare scenarios played continually in my head." (Lukianoff and Haidt 2015b). This strikes me a clear overreaction to a relatively

[33]The two incidents in Haidt's own life that raised his hackles: a student requesting a content warning for a reading about cancer (via a comment on their homework assignment), and another voicing concern (in a student evaluation) about his using a painting of topless women to illustrate the lesson of Odysseus and the Sirens in an introductory psychology lecture. Neither student escalated the complaint and, by Haidt's own admission, they did exactly what he thinks they ought to do in bringing the matter to his attention directly. Moreover, whether or not one agrees with these students' requests, their simply expressing their discomfort strikes me as perfectly reasonable—and par for the course in a healthy pedagogical context. So why didn't Haidt take the note with equanimity? Over-sensitivity is my diagnosis. See Lukianoff and Haidt 2015b.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

small risk: that making an ill-advised or misinterpreted remark in the classroom or another professional setting could lead to protest, punishment, and pariahdom, as opposed to ordinary accountability.

Clearly, the students' supposed sensitivity to harm is not the only one in the offing here. We should be alive to the possibility that, despite the fact that subordinate group members, such as Hari Kondabalu, have typically been the ones accused of over-sensitivity, it may be dominant group members who need to toughen up a bit. In complaining so much about the over-sensitivity of others, they are betraying their *own* over-sensitivity to false accusations, their heightened and hyperbolic sensitivity to harm to white men like them, and their outsized reactions to the relatively small risk of being unfairly "canceled" over a minor misdeed or misinterpretation of their action. Such over-sensitivity is not an innocuous mistake, moreover: it is plausibly being weaponized to keep subordinate group members, like Kondabolu, fearful and silent. When subordinate group members dread being called over-sensitive, it is easy to shut us up and prevent us from protesting injustice. The trope of the "snowflake"—weak and tremulous, on the one hand, yet spoiled and litigious, on the other—has been a remarkably successful weapon in making vulnerable people reluctant to give voice to valid complaints and loudly demand better. If we are to make much-needed social progress, there is a need for moral clarity: sensitivity is valuable, and there are worse things than being slightly over-sensitive to the systemic injustice that is all too often minimized.

Acknowledgments. My thanks to Michelle Kosch, Daniel Manne, Bianca Waked, Willow Starr, students in two of my seminars at Cornell, and audiences at Seattle University, The Ohio State University, Ursinus College, Ithaca College, UC Irvine, and UCLA for their helpful feedback about this paper.

References

Abramson, Kate. 2014. Turning up the lights on gaslighting. *Philosophical Perspectives* 28 (1): 1–30.

Allport, Gordon W. 1954. *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Anderson, Elizabeth. 1995. Feminist epistemology: An interpretation and a defense. *Hypatia* 10 (3): 50–84.

Anderson, Elizabeth. 2020. Feminist epistemology and philosophy of science. In *The Stanford encyclopedia of philosophy* (Spring 2020 ed.), ed. Edward N. Zalta. https://plato.stanford.edu/archives/spr2020/entries/feminism-epistemology

Antony, Louise M. 2000. Situating feminist epistemology. *The Proceedings of the Twentieth World Congress of Philosophy* 8: 31–40.

Bailey, Olivia. 2023. Growing up and getting better: On some moral costs of moral sensibility change, Draft MS.

Banaji, M. R., S. T. Fiske, and D. S. Massey. 2021. Systemic racism: Individuals and interactions, institutions and society. *Cognitive Research* 82 (6). https://doi.org/10.1186/s41235-021-00349-3

Basu, Rima. 2019. What we epistemically owe to each other. *Philosophical Studies* 176 (4): 915–31.

Beauvoir, Simone de. 2011. *The second sex*, trans. Constance Borde and Sheila Malovany-Chevallier. New York: Vintage.

Beeghly, Erin, and Alex Madva, eds. 2020. *An introduction to implicit bias: Knowledge, justice, and the social mind*. New York: Routledge.

Brownstein, Michael, and Jennifer Mather Saul, eds. 2016. *Implicit bias and philosophy*, vol. 1. *Metaphysics and epistemology*. Oxford: Oxford University Press.

Chemaly, Soraya. 2018. *Rage becomes her: The power of women's anger*. New York: Atria Books.

Cherry, Myisha. 2021. *The case for rage: Why anger is essential to anti-racist struggle*. New York: Oxford University Press.

Collins, Patricia Hill. 1990. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. New York: Routledge.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

Cooper, Brittney. 2018. *Eloquent rage: A black feminist discovers her superpower*. New York: St Martin's Press.

Crenshaw, Kimberlé W. 1991. Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review* 43 (6): 1241–99.

Crenshaw, Kimberlé W. 1993. Beyond race and misogyny: Black feminism and 2 Live Crew. In *Words that wound*, ed. Mari J. Matsuda, Charles Lawrence III, Richard Delgado, and Kimberlé Williams Crenshaw. Boulder, CO: Westview Press.

Dastagir, Alia. 2025. *To those who have confused you to be a person: Words as violence and stories of women's resistance online*. New York: Crown.

Dell, Paul F., and John A. O'Neill, eds. 2009. *Dissociation and the dissociative disorders: DSM-V and beyond*. New York: Routledge.

DiAngelo, Robin. 2018. *White fragility: Why it's so hard for white people to talk about racism*. Boston: Beacon Press.

Douglass, Frederick. 1845. American prejudice against color: An address delivered in Cork, Ireland, October 23, 1845. *Cork Examiner*, October 27. In *The Frederick Douglass Papers: Series One—Speeches, Debates, and Interviews*, vol. 1, ed. John Blassingame. New Haven: Yale University Press, 1979.

Dror, Lidal. 2023. Is there an epistemic advantage to being oppressed? *Noûs* 57 (3): 618–40.

Du Bois, W. E. B. 2015. The study of the Negro problems. In *The problem of the color line at the turn of the twentieth century: The essential early essays*, ed. Nahum Dimitri Chandler. New York: Fordham University Press. Reprint of "The study of the Negro problems." *Annals of the American Academy of Political and Social Science* 11 (1) (1989): 1–23.

Du Bois, W. E. B. 2007. *The souls of Black folk*. Oxford: Oxford University Press. Originally published in 1903.

Feeny, Norah C., Lori A. Zoellner, and Edna B. Foa. 2000. Anger, dissociation, and posttraumatic stress disorder among female assault victims. *Journal of Traumatic Stress* 13 (1): 89–100.

Hamilton, Kate (initially "Anonymous"). 2018. We need to talk about sexual assault in marriage. *Vox*, March 8. https://www.vox.com/first-person/2018/3/8/17087628/sexual-assault-marriage-metoo

Hamilton, Kate. 2024. *Mad wife: A memoir*. Boston: Beacon Press.

Harding, Sandra. 1991. *Whose science? Whose knowledge? Thinking from women's lives*. Ithaca, NY: Cornell University.

Haslam, Nick. 2016. Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry* 27 (1): 1–17.

Holroyd, Jules. 2012. Responsibility for implicit bias. *Journal of Social Philosophy* 43 (3): 274–306.

Jorati, Julia. 2023. *Slavery and race: Philosophical debates in the eighteenth century*. New York: Oxford University Press.

Lawrence, Charles R. 1987. The id, the ego, and equal protection: Reckoning with unconscious racism. *Stanford Law Review* 39 (2): 317–88.

Longino, Helen E. 1990. *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton: Princeton University Press.

Lukianoff, Greg, and Jonathan Haidt. 2015a. The coddling of the American mind. *The Atlantic*, September. https://www.theatlantic.com/magazine/archive/2015/09/the-coddling-of-the-american-mind/399356/

Lukianoff, Greg, and Jonathan Haidt. 2015b. The backstory to "The coddling of the American mind." *The Atlantic*, September. https://www.theatlantic.com/magazine/archive/2015/09/greg-lukianoffs-story/399359/

MacKinnon, Catharine A. 1979. *Sexual harassment of working women: A case of sex discrimination*. New Haven: Yale University Press.

Manne, Kate. 2017. *Down girl: The logic of misogyny*. New York: Oxford University Press.

Manne, Kate. 2020. *Entitled: How male privilege hurts women*. New York: Crown.

Manne, Kate. 2023. Moral gaslighting. *Aristotelian Society Supplementary Volume* 97 (1): 122–45.

McClure, Emma, and Regina Rini. 2020. Microaggression: Conceptual and scientific issues. *Philosophy Compass* 15 (4): e12659.

Medina, José. 2012. *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. New York: Oxford University Press.

Mills, Charles W. 2005. Ideal theory as ideology. *Hypatia* 20 (3): 165–84.

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press

Nussbaum, Martha. 2016. *Anger and forgiveness: Resentment, generosity, justice*. New York: Oxford University Press.
Pinker, Steven. 2011. *The better angels of our nature: Why violence has declined*. New York: Viking.
Rogers, Taylor. 2021. Knowing how to feel: Racism, resilience, and affective resistance. *Hypatia* 36 (4): 725–47.
Simeon, Daphne, and Jeffrey Abugel. 2008. *Feeling unreal: Depersonalization disorder and the loss of self*. New York: Oxford University Press.
Srinivasan, Amia. 2018. The aptness of anger. *Journal of Political Philosophy* 26 (2): 123–44.
Steinberg, Marlene, and Maxine Schnall. 2001. *The stranger in the mirror: Dissociation—the hidden epidemic*. New York: Harper Perennial.
Taylor Mill, Harriet. 2022. *Enfranchisement of women*. DigiCat.
Tessman, Lisa. 2005. *Burdened virtues: Virtue ethics for liberatory struggles*. New York: Oxford University Press.
Traister, Rebecca. 2018. *Good and mad: The revolutionary power of women's anger*. New York: Simon & Schuster.
Weekes Schroer, Jeanine, and Lauren Freeman, eds. 2019. *Microaggressions and philosophy*. New York: Taylor & Francis.
Wollstonecraft, Mary. 1792. *A vindication of the rights of woman with strictures on political and moral subjects*. London: Joseph Johnson. Accessed via https://oll.libertyfund.org/titles/wollstonecraft-a-vindication-of-the-rights-of-woman

Kate Manne is a professor at the Sage School of Philosophy at Cornell University. In addition to academic articles in moral, social, and feminist philosophy, she has written three books—*Down girl: The logic of misogyny* (2017), *Entitled: How male privilege hurts women* (2020), and *Unshrinking: How to face fatphobia* (2024). She regularly writes opinion pieces and articles for non-academic audiences, including in the *New York Times*, *The Cut, Washington Post*, *The Atlantic, The Nation,* and *Time*. She also writes the substack newsletter, "More to Hate," about misogyny, fatphobia, culture, politics, philosophy, and more.

Cite this article: Manne K (2026). "Sensitivity and Survival." *Hypatia* 1–22. https://doi.org/10.1017/hyp.2026.10081

https://doi.org/10.1017/hyp.2026.10081 Published online by Cambridge University Press