(What) Are Stereotyping and Discrimination? (What) Do We Want Them to Be?

Alex Madva, California State Polytechnic University, Pomona, alexmadva@gmail.com

––––––––––––––––––––

Erin Beeghly's "Stereotyping as Discrimination" is—characteristically—clear, thorough, and persuasive, rich with incisive arguments and thought-provoking case studies. In defending the view that stereotyping often constitutes discrimination, she makes a powerful case that, "Living ethically means cultivating a certain kind of 'inner' life and avoiding pernicious habits of thought, no matter how culturally pervasive" (Beeghly 2021b, 13). Support for such claims is traced back not just to Aristotle and the Ten Commandments (Beeghly 2021b, 10), but also to critical social traditions. "As feminists and theorists of race have long noted, the most intimate aspects of our selves, including our ways of thinking, agency, and modes of embodiment, are among the mechanisms of group oppression" (Beeghly 2021b, 13).

I offer, first, a tiny friendly amendment to one brief passage, followed by invitations to explore some of the further potential implications of Beeghly's central theses.

Which Counterfactuals When?

Beeghly's analysis of the Epistemic Shortcut Argument explores how "jumping to conclusions" can be discriminatory. She draws on Fanon's report of asking a white person for directions and being answered in pidgin rather than French. Here and elsewhere, Beeghly carefully considers different senses of discrimination that might be at play, and one passage considers the "literal behavioral sense." She writes:

> The idea here is *not* that stereotypic snap judgments constitute behavior per se. Rather, the view is that shortcuts must be contrasted with the relevant alternative, i.e., more laborious and careful ways of judging individuals. More laborious methods of judgment involve doing things in a literal behavioral sense: asking questions, listening carefully, looking up information online, and so forth. Keeping this contrast in mind, here is the proposal. The white train passenger would have taken a longer epistemic route, hence would have *behaved differently*, if a visibly white stranger had approached him. The counterfactual behavior is what grounds the claim that stereotyping constitutes discriminatory treatment (Beeghly 2021b, 5).

I agree with Beeghly that appeals to counterfactuals are necessary in cases like these—although, admittedly, I am less versed in the literature on discrimination and can imagine how these matters are the subject of controversy. I also agree that, in many cases, the relevant counterfactuals will involve either a) judging members of marginalized groups more quickly and carelessly than they would have judged members of privileged groups, or b) judging members of marginalized groups with more scrutiny and even suspicion than they would have judged members of privileged groups (Beeghly 2021b, 6).

But I'm not sure about the relevant counterfactual in this specific case. Had the person asking for directions also been white, would the white passenger really have slowed down to consider things? I suspect the passenger would just automatically speak as they do to most white strangers, in their standard French dialect. They would have relied on a suite of

cognitive shortcuts tailored for public interactions with members of their own race, nationality, etc.

But there are still relevant counterfactuals here, such as those that vary the *discriminator* rather than the *discriminatee*. These include something like a "reasonable person" standard: How would a less racist person have responded, or, to put it in less loaded terms, how would someone who was not stereotyping have acted in these cases? Some of these counterfactuals, with more virtuous question-answerers, will also involve shortcuts, e.g., automatically responding in French to someone who asks you a question in French. But some of them will involve more cognitive labor: listening more carefully, pausing for a moment to take in more information about the situation, thinking about how to respond, and asking follow-up questions before answering.

(What) Are Stereotyping and Discrimination? (What) Do We Want Them to Be?

One of the especially helpful features of Beeghly's paper (and her oeuvre more generally!) is the careful attention to the polysemy of terms like "stereotyping" and "discrimination." Here she considers how stereotyping can count as discrimination across a variety of meanings of "discrimination." These disambiguations are illuminating contributions in and of themselves, and, again, the arguments are compelling.

That said, within this specific paper, Beeghly largely treats these topics in purely "descriptive" and "metaphysical" terms: given precise and neutral definitions of the terms, what are the sufficient conditions for Xs to count as Ys?

There are other questions that we, as critical social philosophers, could ask here. Taking cues from Sally Haslanger's (2012) distinctions between operative, manifest, and target concepts (respectively), we might ask:

> 1. How do people *actually* use the terms/concepts "stereotyping" and "discrimination"?
> 2. How do people *think* and *report* that they use these terms?
> 3. How *should* people use these terms?

Beeghly has explicitly appealed to contrasts between empirically descriptive, metaphysical, and normative/critical approaches in earlier work, such as in her paper arguing against the widespread view that stereotypes are always or typically false. There Beeghly writes:

> Though the essay is arguably an exercise in metaphysics because I evaluate the falsity hypothesis as a thesis about what stereotypes are, it is also an exercise in social philosophy. Part of my argument against the falsity hypothesis will be that it fails to serve the critical function needed for a good ameliorative account of stereotypes (2021a, 33).

In fact, the normative question (3) above can be broken down further, depending on which kinds of "should" we're interested in. Here are two more specific questions:

> 4. How should we use these terms to minimize misunderstanding and facilitate communication?
> 5. How should we use these terms critically, to help us call out and fight injustice and oppression?

Making headway on (4) requires paying attention to (1) and (2). And if we're looking at purely descriptive linguistic/conceptual questions, about how people tend to use these terms, then we have to grapple immediately with the challenge of messy polysemy. These words mean different things in different contexts. The same speakers use these terms in distinctive ways across different situations. More to the point, some communities and subgroups use these terms in one set of ways while other communities use these terms in other ways.

Concretely, there are two particular communities and contexts that seem to treat it as important to distinguish stereotyping from discrimination:

> a. intergroup social-scientific theorizing
> b. anti-discrimination legal practices

(I have more confidence in what I'm about to say about the former than the latter.)

Intergroup Social-Scientific Theorizing

Any textbook you pick up in psychology, sociology, etc., that discusses "prejudice," "stereotyping," and "discrimination" (and any internet search you do of those three terms together) will deliver some version of a familiar and pretty sharp three-part distinction. Here is leading psychologist Susan Fiske's (2021) recent gloss:

> People are often biased against others outside of their own social group, showing prejudice (emotional bias), stereotypes (cognitive bias), and discrimination (behavioral bias).

> … people put other people into groups, using that label to inform their evaluation of the person as a whole—a process that can result in serious consequences. This module focuses on biases against social groups, which social psychologists sort into emotional prejudices, mental stereotypes, and behavioral discrimination. These three aspects of bias are related, but they each can occur separately from the others (Dovidio and Gaertner, 2010; Fiske, 1998). For example, sometimes people have a negative, emotional reaction to a social group (prejudice) without knowing even the most superficial reasons to dislike them (stereotypes) …

> Prejudice is an evaluation or emotion toward people merely based on their group membership.
>
> Stereotype is a belief that characterizes people based merely on their group membership.
>
> Discrimination is behavior that advantages or disadvantages people merely based on their group membership.

Beeghly is, of course, well aware of these well-worn social-scientific definitions (e.g., Beeghly 2015). But to get at long last to the punchline, even though Beeghly is right that stereotyping sometimes or always constitutes discrimination, there are further questions to ask about potential payoffs and perils of actually *speaking that way*. For example, given that social scientists explicitly cordon off discrimination (*qua* behavioral stuff) from stereotyping (*qua* cognitive stuff), taking Beeghly's proposals on board might risk miscommunication and misunderstanding, setting back efforts for interdisciplinary crosstalk and collaboration. Some of my questions here are:

> • Should we push for using these terms interchangeably in many of the contexts where stereotyping constitutes discrimination? What would it actually look like if we started speaking as if stereotyping was often, just as such, discrimination? And that wrongful stereotyping was just as such wrongful discrimination?
>
> • Or should we, in general or in certain specific contexts, continue to reserve the term "stereotyping" for the more "purely cognitive" stuff (or the *ostensibly* or *comparatively* more cognitive stuff—like Beeghly's example of stereotyping fellow bus-riders while mind-wandering and staring out the window)?
>
> • Finally, would Beeghly advocate that social scientists rewrite their textbooks? If so, how much revision is in order? Is it still OK for them to list the three terms, as long as they immediately follow them up with clarifications that reflect Beeghly's compelling points?
>
> • Or should we trash the tripartite distinction altogether? What remaining purpose, if any, does it serve?

These questions are entangled with broadly empirical and open-ended issues, probably not best settled from the armchair, but I'm curious what Beeghly thinks about them.

Readers might infer from the tenor of this discussion that I am advocating for preserving the standard terminological classifications. Shame on you for taking epistemic shortcuts! I am already on the record as challenging the "prejudice" vs. "stereotype" distinction, especially with respect to *implicit* stereotypes and prejudices (Madva and Brownstein 2018). Although the *conceptual* distinction between "evaluations/emotions" on the one hand and "beliefs" on

the other is intuitively clear and familiar enough, the *psychological* questions about whether there are two genuine mental kinds here, and if so, how to distinguish them, remain up in the air. Michael Brownstein and I argued, for example, that all implicit stereotypes are evaluative and affect-laden.

Subsequently, social psychologists have subjected the prevailing prejudice-stereotyping distinction to empirical scrutiny (Kurdi et al. 2019; Phills, Hahn, and Gawronski 2020). Their work casts doubt on popular theories that treat implicit prejudices and stereotypes as grounded in distinct neural substrates (e.g., Amodio and Ratner 2011). It calls into question intuitive but testable claims like Fiske's that people can "have a negative, emotional reaction to a social group (prejudice) without knowing even the most superficial reasons to dislike them (stereotypes)." This research also challenges converse hypotheses, e.g., that people can harbor "mere beliefs" about social groups devoid of evaluative or emotional significance. On one interpretation of the relation between prejudices and stereotypes, for example, a person's evaluative (prejudicial) responses to a person simply reflect the net valence of all their momentary (stereotypical) thoughts about them (Gawronski and Bodenhausen 2006).

In a similar vein, Beeghly's (2021b, 6–7) appeals to embodied cognition suggest that we should be skeptical that people can have much in the way of "mere social thoughts and feelings" that never express themselves in behavior in problematic ways. The truth will out, and so will people's pernicious ways of thinking and feeling about social groups and their members. As she notes, the point is not just that our inner lives inevitably "leak out," but that the outer is partially constitutive of the inner. Vast swaths of our social thoughts and feelings may have essentially motivational and action-oriented features, guiding our goals and motor orientation to the world.

So, probably, I would urge Beeghly to go ahead and advocate for substantive revisions to social-scientific dogmas and textbooks. One less radical way to do this would be to treat thoughts, feelings, and behaviors as three different aspects or components of a single social-cognitive construct, rather than as three self-standing constructs of their own (see also Brownstein and Madva 2012b; 2012a; Gendler 2008). There is, interestingly enough, already precedence for this sort of approach within social psychology. Although "prejudice" is often defined as a purely affective social response, it is also often defined as a kind of "attitude," which (as social psychologists use the term) is comprised of three components. Can you guess what three they are?

> Prejudice is typically conceptualized as an attitude that, like other attitudes, has a cognitive component (e.g., beliefs about a target group), an affective component (e.g., dislike), and a conative component (e.g., a behavioral predisposition to behave negatively toward the target group) (Dovidio et al. 2010, 5).

Theorizing thoughts, feelings, and actions as a "package deal" in this way might represent a more promising way forward. In any case, the fact that prejudice is so variously and

circularly defined is a good indicator, if anything is, that there's ample room for clarifying and revising prevailing social-scientific discourse and theory.

Yet I still have some reservations about leaping headlong into the full-blown conflation of "stereotyping" and "discrimination." What would it mean, I wonder, for anti-discrimination legal practitioners to linguistically and conceptually reengineer their practices along Beeghly's lines?

Anti-Discrimination Legal Practices

It goes without saying that some of the most important avenues for resisting injustice and oppression go through the courts and anti-discrimination law. I am less knowledgeable about the parlances of those practicing anti-discrimination law than of those practicing social science. Some of what I know is based on Beeghly's other work, as when she notes the disagreements among Supreme Court Justices regarding the meaning and nature of stereotypes (Beeghly 2021a, 34–35). But it seems that some version of the tripartite distinction is operative in legal contexts as well. Legal action around discrimination targets how people are actually treated, whether by other individuals in formal, informal, or unstructured settings, or by those crafting or acting in accordance with facially or indirectly discriminatory laws, norms, and regulations—as opposed to targeting how people "merely" think and feel.

Again, I think Beeghly is conceptually and metaphysically correct that stereotyping can be discriminatory. Acknowledging this, the key question is whether we might still want, for our multifarious legal purposes, to continue granting that there can be a lot of daylight between stereotyping (merely thinking bad thoughts) and discrimination (committing wrongful actions and writing unjust laws, sometimes on the basis of those bad thoughts). If we start thinking *and talking* seriously about wrongful forms of stereotyping not just as morally condemnable and in need of amelioration but as full-blown discrimination, does or should anything follow legally speaking? It's one thing to acknowledge, as I agree with Beeghly that we should, the moral imperative of cultivating more virtuous inner lives. But do the practical upshots explored by Beeghly extend beyond the ethical realm and into the legal? If so, how? If not, why not?

If it's true that in legal settings, part of *what it means* to call something discriminatory is to say that it is or should be legally actionable, does it follow that categorizing stereotyping as discrimination can make stereotyping somehow legally actionable, too? The elephant in the room can be named in two words: thought police. Pursuing this approach would likely raise all sorts of hackles about the Big Data Big State monitoring and regulating not just our speech and action but our very thoughts. There will, of course, be tons to say here, from observations of the presumably myriad objectionable and unobjectionable ways in which the government and other powerful corporate actors already try to shape our thoughts (with varying degrees of success) to questions about how one could even measure or prove what people were thinking independently of observing their behavior, and so on.

One way to preserve something like the three-part distinction for legally actionable purposes might be to treat each category as best-suited for picking out a specific *subset* of "behavioral" biases. Economist Marie-Anne Valfort (2018), for example, suggests that we can think of prejudice as "taste-based discrimination" (e.g., when you don't hire someone because you just don't like members of that group) and stereotyping as "statistical discrimination" (e.g., when you don't hire someone because you believe, whether on accurate or inaccurate grounds, that members of their group will take more sick days than members of other groups). I don't think Beeghly would buy into Valfort's particular way of sorting things, but there might be some traction to the idea that stereotyping represents one important subset of discrimination, subject to a somewhat distinctive set of norms and modes of practical, moral, and legal redress, in contrast to other subsets of discrimination.

Symmetry and Asymmetry Between Belief and Action

I'm very sympathetic with Beeghly's proposal that many of the wrong-making features of behavioral discrimination will also apply to "mere" in-the-head stereotyping. That said, the strongest argument for normative asymmetry (or, more modestly, the strongest argument against terminological conflation) here might be something like the following: normatively speaking, whether we are crafting laws and policies or deciding how to act in particular cases, we want our planning to be based on as accurate a view of the world as possible. As Beeghly has long argued (e.g., 2015; 2021a), stereotypes have various important roles to play in giving us an accurate picture of the world. It's important for us to be aware, first, of the prevailing pernicious stereotypes that shape people's interpretations of and reactions toward others and themselves. But, second, if we're defining stereotypes very generally to include "beliefs about typical group traits" (Beeghly 2021b, 2), then that means that even generalizations like "Black drivers are more likely to be pulled over by the police" and "Black children are more likely to be raised in single-parent homes," are stereotypes, even though my belief in them is both true and justified. These are important facts that we *should* attend to in many deliberative contexts (Madva 2016). These generalizations often reflect, in sometimes straightforward but often complex ways, injustices we are called to resist.

To the extent that such stereotypes have important roles to play in giving us an accurate picture of the world, we shouldn't try to "ban them from thought" so much as we should try to take care in "putting them into action." Calling these thoughts discriminatory—even if we are at pains to clarify when we mean "discriminatory" in a neutral versus evaluative sense— might impede our ability to confront the social facts as they really are in order to decide how we want them to be.

Alternatively, the foregoing might not be so much a case for normative asymmetry as one gesture toward fleshing out Beeghly's insistence on pluralistic and context-dependent approaches to these matters. She rightly notes that one can defend normative symmetry "while remaining attentive to the special normative features of thought, speech, actions, and policy and without papering over important ethical distinctions" (2021b, 13).

Or maybe much of the huffing and puffing of this commentary boils down to vague concerns about the limits of our collective linguistic agency. No matter how much academics and activists try to get folks to think of "racism" as primarily a structural rather than an individualistic wrong, public conversation can't seem to escape the gravitational pull of the individualistic interpretation. Similarly, to the extent that the folk, and the social scientists, and the lawyers, harbor such ingrained, preconceived notions of both the *meaning* and the *intrinsically negative valence* of "stereotyping" and "discrimination," is trying to push back on the prevailing ways of speaking about these ideas like trying to plow the sea? I hope not, but in the short term, our time might be better spent working for social change while accommodating ourselves to these popular ways of speaking, even as Beeghly compels us to appreciate just how confused and half-baked the prevailing preconceptions are.

References

Amodio, David M., and Kyle G. Ratner. 2011. "A Memory Systems Model of Implicit Social Cognition." *Current Directions in Psychological Science* 20 (3): 143–48.

Beeghly, Erin. 2021a. "What's Wrong with Stereotypes? The Falsity Hypothesis." *Social Theory & Practice* 47 (1): 33–61.

Beeghly, Erin. 2021b. "Stereotyping as Discrimination: Why Thoughts Can Be Discriminatory." *Social Epistemology* 35 (6): 547–63. Beeghly, Erin. 2015. "What Is a Stereotype? What Is Stereotyping?" *Hypatia* 30 (4): 675–91.

Brownstein, Michael, and Alex Madva. 2012a. "Ethical Automaticity." *Philosophy of the Social Sciences* 42 (1): 68–98.

Brownstein, Michael, and Alex Madva. 2012b. "The Normativity of Automaticity." *Mind & Language* 27 (4): 410–34.

Dovidio, John F., Miles Hewstone, Peter Glick, and Victoria M. Esses. 2010. "Prejudice, Stereotyping and Discrimination: Theoretical and Empirical Overview." In *The SAGE Handbook of Prejudice, Stereotyping and Discrimination*, edited by John F. Dovidio, Miles Hewstone, Peter Glick, and Victoria M. Esses, 3–28. SAGE Publications.

Fiske, Susan T. 2021. "Prejudice, Discrimination, and Stereotyping." In *Noba Textbook Series: Psychology*, edited by Robert Biswas-Diener and Ed Diener. Champaign, Illinois: DEF Publishers.

Gawronski, Bertram, and Galen V. Bodenhausen. 2006. "Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change." *Psychological Bulletin* 132 (5): 692–731.

Gendler, Tamar Szabó. 2008. "Alief in Action (and Reaction)." *Mind & Language* 23 (5): 552–85.

Haslanger, Sally. 2012. *Resisting Reality: Social Construction and Social Critique*. OUP USA.

Kurdi, Benedek, Thomas C. Mann, Tessa E. S. Charlesworth, and Mahzarin R. Banaji. 2019. "The Relationship between Implicit Intergroup Attitudes and Beliefs." *Proceedings of the National Academy of Sciences*, March, 201820240.

Madva, Alex. 2016. "Virtue, Social Knowledge, and Implicit Bias." In *Implicit Bias and Philosophy: Metaphysics and Epistemology: Volume 1*, edited by Michael Brownstein and Jennifer Saul, 191–215. Oxford: Oxford University Press.

Madva, Alex and Michael Brownstein. 2018. "Stereotypes, Prejudice, and the Taxonomy of the Implicit Social Mind." *Noûs* 52 (3): 611–44.

Phills, Curtis E., Adam Hahn, and Bertram Gawronski. 2020. "The Bidirectional Causal Relation Between Implicit Stereotypes and Implicit Prejudice." *Personality and Social Psychology Bulletin*, February, 014616721989923.

Valfort, Marie-Anne. 2018. "Do Anti-Discrimination Policies Work?" *IZA World of Labor*, May.