Shadowboxing with Social Justice Warriors

A Review of Endre Begby's Prejudice: A Study in Non-Ideal Epistemology, Oxford University Press, New York, 2021, 240 pp. $40.00 (hardback), ISBN: 0198852834

Alex Madva

Published online: 05 Oct 2022.

Submit your article to this journal ⬚

Article views: 73

View related articles ⬚

View Crossmark data ⬚

Routledge
Taylor & Francis Group

BOOK REVIEW

Shadowboxing with Social Justice Warriors

A Review of Endre Begby's Prejudice: A Study in Non-Ideal Epistemology,
Oxford University Press, New York, 2021, 240 pp. $40.00 (hardback), ISBN:
0198852834

Endre Begby's *Prejudice: A Study in Non-Ideal Epistemology* engages a wide range of issues of enduring interest to epistemologists, applied ethicists, and anyone concerned with how knowledge and justice intersect. Topics include stereotypes and generics, evidence and epistemic justification, epistemic injustice, ethical-epistemic dilemmas, moral encroachment, and the relations between blame and accountability. Begby applies his views about these topics to an equally wide range of pressing social questions, such as conspiracy theories, misinformation, algorithmic bias, discrimination, and criminal justice.

Through it all, the book's central thesis is that prejudices can be epistemically rational, a corrective against what Begby takes to be the received view that prejudices are always and everywhere bad:

> when philosophers attempt to view prejudice through the normative lens of epistemology, their conclusion is typically swift, clear, and likewise negative: no one can be epistemically justified in holding a prejudiced belief; prejudiced belief, whenever it occurs, is a symptom of some kind of breakdown of epistemic rationality. *(pp. 1–2)*

The first three chapters define the relevant terms and articulate Begby's background epistemic assumptions and methodology. Chapter 4 then broaches prejudice acquisition, where the crux of Begby's argument is that when, say, the best students in Johnny's math class are boys, it may be inductively rational for Johnny to infer that boys are better at math than girls. Begby also notes that many prejudices are acquired through testimony, and it is often rational to believe what others tell us. Chapter 5 considers prejudice maintenance, pointing out that it may be rational for us to hold onto our prejudices in the face of counterexamples. Encountering one mathematically talented girl, for example, is consistent with boys being *on average* better at math, especially if Johnny has reason to think the girl he encounters is not representative of girls in general.

Chapters 6 through 8 then "provide significantly more detail by outlining specific social and institutional mechanisms by which prejudiced beliefs might be boosted, disseminated, or insulated from contrary evidence" (p. 95). These include ways of conveying prejudiced beliefs that prevent the recipient of one's testimony from taking future counterevidence seriously (Chapter 6), ways that prejudice can structure our interactions through widely shared social scripts even when individuals don't believe

the prejudices are true (Chapter 7), and ways that algorithms can perpetuate prejudice when fed data from "a deeply unequal and unjust society" (Chapter 8, p.135). The book concludes with several chapters exploring the interfaces between ethics and epistemology, examining whether ethics and epistemology might conflict (also Chapter 8), whether morality might shape what we epistemically ought to believe (Chapter 9), and whether individuals can still be held morally responsible when they endorse prejudiced beliefs that are supported by their available evidence (Chapter 10). Importantly, this concluding chapter seeks "a viable platform for vindicating the rights of victims of prejudice and for considering prejudice as a problem that might warrant political intervention" (pp.184–5), even in cases when the prejudices in question are epistemically rational.

Evaluating Begby's organizing claim that prejudice can be rational requires getting clear on his terminology, which is so idiosyncratic as to render the book's title, and much of its content, misleading. As Jessie Munton (2022) notes, the book is not really about *prejudice*, which is widely understood to be an objectionable, affect-laden, favorable (positive) or unfavorable (negative) social attitude. Begby, however, drops the objectionable and affect-laden overtones of the term and opts to use "prejudice" to mean "prejudiced beliefs," or *stereotypes* (1.1). He then restricts his focus to "negatively charged" stereotypes, where the "negativity" refers exclusively to the valence of the stereotype (thinking unfavorably about a group of people) rather than to its epistemic or moral status. These terminological choices make the book difficult to evaluate and situate in contemporary debates. By analogy, consider calling a book *Problematic Emotions* and then stipulating that you're talking about a specific subset of *negative "emotional beliefs,"* and then arguing that, given this idiosyncratic stipulation, such beliefs are not (*necessarily*) problematic. One could of course *defend*, e.g., the view that what we call "prejudices" are really just like unproblematic emotions, or a cognitivist theory of emotions as a species of belief-like state, but readers would expect such claims to depend on arguments and, preferably, on appeals to some kind of quantitative or qualitative evidence, history, memoir, etc.—rather than on terminological stipulation.

Similarly, as Robin McKenna (2022) explains, the book is not an exercise in non-ideal philosophy. Instead of pursuing the "non-ideal" methodologies one finds in, say, critical social (feminist, antiracist, etc.) philosophies, Begby uses "non-ideal" where epistemologists from W.V.O. Quine to Louise Antony would say "naturalized," i.e., pursuing a theory of knowledge based on realistic appraisals of human cognitive limitations and informational ecologies. Non-ideal and naturalist approaches are not mutually exclusive, and they share a commitment to grounding philosophical analysis in accurate accounts of the way the world works, but they are not the same.

Whatever we call Begby's methodology, his self-imposed mandate to take minds as they are is in tension with his unexplained decisions about which topics require careful attention to empirical literature, which can be brushed aside with a footnote, and which can be settled from the armchair. When Begby explains why the book will *not* discuss implicit bias (2.3), he offers a detailed discussion of nuanced psychometric controversies surrounding the Implicit Association Test related to test-retest reliability, effect sizes, and meta-analyses about predictive validity. By contrast, no empirical appeal is evidently needed when it comes to asserting that stereotypes are generic beliefs, or that there is, distinct from belief, an important psychological category called "acceptance." (For Begby, acceptance is "an executive decision to treat a proposition as sufficiently evidenced to act upon in a particular context" (p. 161).) But what is the test-retest reliability of self-reported generic beliefs? Where are the meta-analyses demonstrating that "acceptances" both dissociate from beliefs and predict behavior? To clarify,

philosophical and empirical research into generics and acceptances is fascinating and important; my bafflement regards what Begby takes to entitle his exceedingly strong yet entirely speculative claims about them, in light of his non-ideal/naturalist commitments. For example, Begby (2.1) dismisses the relevance to his project of the entirety of Lee Jussim and colleagues' decades of research on stereotype accuracy by asserting that (apparently *all*) stereotypes are generics, whereas Jussim finds that at least some stereotypes reflect alternative cognitive-propositional structures and statistical relationships. Begby also writes as if there is only one *kind* of generic, with but one evidence-tracking and inferential profile. He offers no evidence to support such claims.

The felt absence of empirical references is most palpable in Begby's discussion in Chapter 6 of "evidential preemption," a phenomenon captured by statements such as, "My opponents will tell you that p; but I say q" (p. 96). Evidential preemption aims to "'inoculate' recipients against future contrary evidence" (p. 95). Begby's analysis neglects 60 years of research on this phenomenon under a slightly different heading, *inoculation theory*, and, more recently, *prebunking* and *refutational preemption*. Thus McGuire (1961, p. 326) experimentally demonstrates that "beliefs can be 'inoculated' against persuasion in subsequent situations … by pre-exposing the person to the counterarguments in a weakened form." The relevant literature is vast. A quick search for "inoculation theory" generates over a million hits on Google and over 3,400 hits on Google Scholar. The topic has its own Wikipedia page and countless explainers on YouTube. Begby mentions none of it.

Notably, this literature is long past documenting, analyzing, and taxonomizing the phenomenon and has moved on to *harnessing* inoculation strategies to prevent rather than propagate misinformation. One outcome of this research is the entertaining online game, *Bad News*, which launched in February 2018. Several studies confirm that playing for just 15 minutes delivers a "psychological inoculation against not only specific instances of fake news, but the very strategies used in its production" (Basol et al., 2020, p. 1). The extensive evidence behind inoculation theory renders Begby's speculations – about how evidential preemption *might* be relevant to understanding and combating conspiracy theories, echo chambers, and fake news (6.3.1)—superfluous. One might as well coin a new term for "nudges" (let's say "prods") and then, without mentioning the vast existing literature, speculate that prods to auto-enroll in savings plans could help people prepare for retirement.

This empirical inattention persists through the book's central arguments that prejudice can be rational. Again, it is fair game to argue that many of what we ordinarily think of as pernicious prejudices and stereotypes are rational. Yet despite his emphasis on non-ideality, Begby's arguments are overwhelmingly supported by idealized depictions of hypothetical characters, rather than engagement with the voluminous, complex, fascinating, and highly contested historical, autobiographical, and scientific literatures exploring the psychological and structural factors driving prejudiced belief (e.g., in developmental, political, cognitive, or intergroup psychology). What's more, all the protagonists in Begby's cases clearly code as masculine, white, European or American (or, in one case, Saudi; p.93) youth, e.g., Harry forming "prejudices" about Slytherin based on Ron's testimony (4.2.1). In nearly every case, a white masculine paragon of intellectual virtue operates in circumstances of extreme evidential insulation, and as a result forms negative stereotypes about women or people of color. Uncontroversially, the more insulated people are from the truth, the more intuitive it will seem that they are rational in acquiring and maintaining prejudiced or false beliefs. (Radical versions of epistemic insulation like this are just skeptical scenarios with Evil Demons or envatted brains.) Begby assures readers that his idealizations reflect

"plausible epistemic contexts" (pp. 70, 93), but he makes no effort to assess how common or realistic they are. For example, Chapter 4 on prejudice acquisition cites none of the empirical literature on child development, attitude learning, conditioning, or belief formation – unless we think the relevant empirical work began and ended with Gordon Allport (1954)! In fact, Begby repeatedly, explicitly ducks difficult questions about the flesh-and-blood epistemic-psychological dispositions and informational contexts of real people, to instead argue that some prejudiced people *might* be rationally responding to evidence:

> None of this is to deny that many instances of prejudice may (also) be held in place by the irrational contrivances of bias, motivated believing, affected ignorance, or similar mechanisms. It is, rather, to deny that they must *always* or *necessarily* involve some such distinctive and specifiable manner of epistemic irrationality. *(p. 94)*

Begby's focus on the sheer possibility of rational prejudice is inconsonant with his overarching goals to describe and understand current social realities. This focus on sheer possibility is tough to square, for example, with his claim midway through the book that his method has been to "observe" how prejudice "tends to spread from person to person by mechanisms of overt communication" (p. 114).

And who today would deny that negative stereotypes "could ever be justified" (p. 114)? Begby provides almost no compelling examples of serious contemporary work endorsing what he calls "the common picture" (pp. 73, 84), i.e., that prejudices (in Begby's idiosyncratic sense) *are always necessarily* irrational. Without clear interlocutors, Begby shadowboxes with straw arguments, which he must read between others' lines to find. For example, he takes objection to a claim that Quassim Cassam "doesn't assert, but certainly seems to imply" (p. 71), to the effect that only first-hand experience (rather than testimony) could justify the prejudiced belief that all politicians lie. But maybe Cassam doesn't assert it because he doesn't believe it? In a subsequent footnote, Begby disagrees with Cassam's claim that belief in conspiracy theories flows from intellectual character flaws. After all, "there may be nothing wrong, per se, with the intellectual character of" conspiracy theorists (p.110 n.32). Yes, there *may* be nothing wrong with some conspiracy theorists' intellectual dispositions; Cassam simply suggests (as against situationists) that "it is *sometimes* appropriate to explain questionable beliefs by reference to intellectual character vices," like gullibility, dogmatism, closed-mindedness, and negligence (Cassam, 2016, p. 159 emphasis added). Unless Cassam argues that it is *impossible* for someone to succumb to prejudiced beliefs and bogus conspiracy theories following epistemically impeccable procedures, there is no disagreement. Begby provides no such textual evidence, nor can I find any, and it's hard to envision why Cassam or anyone else would say something so strong. (I do not mean to suggest that Begby *only* defends uncontroversial claims against straw opponents. Upfront, he endorses an internalist and evidentialist epistemology, which puts him in the thick of contentious ongoing debates. It's a further question whether Begby's arguments for these views make substantive contributions – versus insisting flatfootedly on the familiar intuitions, cases, and arguments that his opponents are already aware of and responding to.)

Indeed, there's a long list of would-be interlocutors whom Begby treats uncharitably, especially when it comes to putative endorsements of some variant of the view that prejudice (in this idiosyncratic sense) is always everywhere irrational. One recurring pseudo-opponent is Miranda Fricker. In Chapter 5, Begby cites Fricker along with Nomy Arpaly, Michele Moody-Adams, and others for the argument that some social

beliefs which are rationally acquired may nevertheless prove to be mere prejudices when individuals cling to them in the face of counterevidence. Begby thinks it's an objection to this claim to construct hypothetical cases (and massage others' hypothetical cases) such that prejudiced people don't confront *enough* counterevidence to upend their priors. Again uncontroversially, if we stipulate that you have rationally acquired the generic belief that "white men can't jump," but then you bump into one white guy with an eye-popping vertical leap, that's not yet enough to overturn your belief. As is well known in social philosophy, linguistics, etc., universal generalizations don't admit exceptions, but generics do. While Chapter 5 walks through variants of cases like this, the reader wonders where and why anyone would disagree.

With straw enemies like these, who needs friends? More perplexing than Begby's portrayal of ostensible opponents is his engagement with, and contributions relative to, the work of alleged allies. The book is littered with passages that effectively say "My claim is *p*," followed by footnotes like, "For a similar argument for *p*, see so-and-so." The cursory nature of these references – many of which happen to be to women, people of color, junior scholars, and numerous writers in feminist, antiracist, and continental traditions – leaves the reader with the impression that they were absent from Begby's earlier manuscripts, and added perfunctorily in response to feedback. As these see-similar-existing-argument references rack up, however, the question becomes increasingly pressing just what this monograph's distinctive contributions are meant to be. Readers familiar with this literature will know that many arguments for something like the "rationality of prejudice" have already been explored in insightful philosophical depth. Feminist philosophers going back centuries, at least to, say, Mary Wollstonecraft (1792), John Stuart Mill and Harriet Taylor Mill (1869), etc., have argued that one of the great difficulties of combating gender oppression is the surface-level evidence supporting sexist beliefs. Begby cites a subset of the contemporary theorists (there are too many to list here) developing these ideas in various sophisticated ways but, because he doesn't clarify how his views relate to or build on theirs, their contributions, as much as his, are obscured. What to say about the irony that a book about prejudice systematically minimizes and misreads the philosophical insights and traditions of members of marginalized groups? That a book marked by tremendous compassion for people living in insulated "communities where ... evidence [that prejudices are false] is systematically suppressed" is itself insulated from, suppressing, and appropriating influential work on these topics?

Thus the final questions to raise regard the book's intended readership. Given Begby's aim to pick apart various "*very* naïve" views (44, original emphasis), one might suppose the text is geared toward beginners, e.g., undergraduates, laypeople, or academics not already immersed in the relevant debates. But until roughly Chapters 8 through 10, Begby does not consistently explain technical terms and constructs. In one paragraph (p. 46), for example, he contrasts our human cognitive limitations with "Leibniz's God" and "Laplace's Demon" without explaining who these gods and demons are, or pointing readers to a citation of Laplace or anyone who discusses his work. Most strikingly, Begby's introduction of generics (5.2.1)–a concept foundational to his account of prejudice and his argument for its rationality – runs a mere two paragraphs and does not actually explain what generics are. The book is manifestly written for people who already know how propositions like "ducks lay eggs" differ from other propositional structures.

The upshot is that it's not clear who this book is for. If it's too advanced for students, too rehashed and uncharitable for those actively working on these areas, and too speculative for those seeking a naturalistic, non-ideal account of social reality, then

perhaps it's for philosophers who share a "broadly progressive political mindset" (p. 141) but are concerned about the irrational excesses and empathy deficits of the social-justice left? Perhaps the book is written for people who would casually take up and repeat Allport's (1954) archaic reference to Asian people as "Orientals" (p. 66), and folks who might enjoy a little joke about being reluctant to read your children books with made-up titles like "*Imbecilic Indians* or *Krazy Koreans*" (p. 164, n. 15). Or perhaps the book is for all the Harrys and Johnnys of the world who don't like seeing their friends or family members (or themselves) getting called names on the internet. Although the book culminates in the argument that prejudices – even pristinely rational ones – may still violate their targets' rights and thereby warrant compensation, Begby nevertheless insists that *blaming* prejudiced people is inappropriate. Apparently, what the world needs now is another plea on behalf of racist, sexist white men everywhere: Forgive them, Social Justice Warriors, for they know not why they're wrong . . . at least not always and necessarily.

References

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

Basol, M., Roozenbeek, J., & Van der Linden, S. (2020). Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of Cognition*, *3*(1), 2. https://doi.org/10.5334/joc.91

Cassam, Q. (2016). Vice epistemology. *The Monist*, *99*(2), 159–180. https://doi.org/10.1093/monist/onv034

McGuire, W. J. (1961). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. *Journal of Abnormal and Social Psychology*, *63*(2), 326–332. https://doi.org/10.1037/h0048344

McKenna, R. (2022). *Review of* Prejudice: A Study in Non-Ideal Epistemology. https://ndpr.nd.edu/reviews/prejudice-a-study-in-non-ideal-epistemology/

Mill, J. S., & Mill, H. T. (1869). *The subjection of women*. Longmans, Green, Reader, and Dyer.

Munton, J. (2022). Prejudice: A study in non-ideal epistemology: By Endre Begby. *The Philosophical Quarterly*, *72*(4), qac006. https://doi.org/10.1093/pq/pqac006

Wollstonecraft, M. (1792). *A vindication of the rights of woman: with strictures on political and moral subjects*. J. Johnson.

Alex Madva

Department of Philosophy, California State Polytechnic University, Pomona

✉ alexmadva@gmail.com

Check for updates