

REPLIES TO CONTRIBUTORS

Katherine Puddifoot¹ 

¹ Durham University, United Kingdom

Discussion – Received: 01/11/2024 Accepted: 25/01/2025

This paper is part of a book symposium on Katherine Puddifoot's *How Stereotypes Deceive Us* guest edited by Marina Trakas (IFF, CONICET)

ABSTRACT

This paper provides responses to the 4 commentaries by Federico José Arena, Leonie Smith, Federico Picinali, and Jennifer Saul under the main headings: “Definition of stereotypes”; “Single/dual factor view”, “Epistemic benefits of egalitarian beliefs”, “Beyond stereotyping beliefs”, “Which disposition?”, “More radical implications of evaluative dispositionalism”, “Stereotypes, reality and testimonial injustice”, “Normative stereotypes”, and finally “Moral encroachment”.

Keywords: epistemic benefits; dispositions; stereotypes; single/dual factor view.

Introduction

How Stereotypes Deceives Us (HSDU) primarily aims to provide a characterisation of the ways that stereotypes lead to misperceptions and misunderstandings of people and events. It argues that stereotypes can have this negative impact in various ways, including when a stereotype reflects an aspect of social reality. I defend a multifactorial approach to stereotyping, according to which multiple factors determine whether any act of stereotyping increases or decrease the likelihood of a misunderstanding or misperception. I develop a view that I call *evaluative dispositionalism*, which says that given the multiple ways that stereotypes can deceive us, we ought to consider, when evaluating any act of believing a stereotype, both the dispositions displayed when acquiring the belief and the dispositions that a person acquired by believing. I articulate implications of the multifactorial view and evaluative dispositionalism for when stereotyping is ethically wrongful, the ways that medical decision-making should be conducted, and for how we should approach any decision about whether to disclose information relating to a stigmatized social identity.

HSDU attempts to integrate a wide-ranging literature, drawing from sources not based on their disciplinary background but instead their quality and relevance to the topics under discussion. Given the breadth of the topic, I could not hope to do justice to, or integrate insights from, all relevant literature, but my hope is that the book might stimulate further discussion—within my home discipline philosophy, but also perhaps more broadly—about the nature and potential epistemic pitfalls associated with stereotyping. This symposium is therefore pleasing because the contributions represent careful and considered responses by experts from inside philosophy and beyond. Below I outline responses by theme rather than by author because in some cases the arguments presented by contributors dovetailed. I aim to show how my ideas have developed in response to symposium contributions.

1. Definition of stereotypes

Let us begin by focusing on the definition of stereotypes. In HSDU I provide a defence of a non-normative conception of stereotypes and stereotyping, that does not define stereotypes as false or stereotyping as misleading (for other defences of a non-normative approach, see Beeghly 2015; Fricker 2007; Jussim et al. 2012; Kahneman 2011). My reason for adopting this approach is pragmatic: to define stereotyping in a way that emphasises the continuity between the mental states involved

with stereotyping and other cognitive states, rather than assuming a discontinuity, with stereotypes separated from other biases like heuristics on the basis that they are always false and misleading (I am influenced here by Ashmore and Del Boca 1981 and Beeghly 2015). In addition to this, I argue that stereotypes are necessarily comparative: they involve a comparison between social groups, suggesting that members of one social group are more likely than some others to possess a particular trait or traits. Here is the definition of stereotypes:

Social attitudes that associate members of some groups more strongly than others with a certain trait or traits. (HSDU, 13)

Saul raises a potential counterexample to this definition. Suppose someone associates people who teach in Durham, a city in the North of England, more strongly than others with teaching in the North of England. Saul suggests that this association could count as a stereotype on my definition, but that this doesn't seem right. I agree that an association between people living in North of England and teaching in Durham does not seem at face value to be a stereotype. So, how should I respond?

There are at least three general options available to me. The first is *bullet biting*. Although it might seem strange to classify this as an example of stereotyping, people's intuitions about what counts as a stereotype are inconsistent. As I discuss in the book, some people seem to intuitively endorse a normative account of stereotyping, according to which stereotypes are always false and misleading (cf. Blum 2004). Others seem to be open to saying that stereotyping can be useful, and necessary, because stereotypes can be accurate heuristics (cf. Beeghly 2015; Fricker 2007; Jussim et al. 2012; Kahneman 2011). Any specific definition of stereotypes is not going to satisfy everyone's intuitive judgements, and so will require some adjustments to classificatory practices. It might be that one appropriate adjustment to classificatory practices is to accept that attitudes like the one associating living in the North of England with teaching at Durham can be stereotypes.

A second option would be to maintain the current definition, but stipulate that definitions or definition-like propositions cannot be stereotypes, for example:

Social attitudes that associate members of some groups more strongly than others with a certain trait or traits *but are not definitions or definition-like*.

Why might one make this move? The thought is as follows. It might seem that one is not stereotyping when more strongly associating people who teach in Durham with working in the North of England because working in Durham, i.e. that characteristic that makes them a member of the target group, by definition involves having the attribute that is ascribed by the stereotype, i.e. working in the North of England. Or, more precisely, it might be said that the characteristic that makes the person a member of the target group almost by definition involves having the attribute ascribed because there may be some exceptional cases, such as people who work in Durham but wholly online, who have never been to Durham, and so forth. If one were to say that people who teach in Durham work in the North of England one would be saying something that is almost true by definition. By stipulating that propositions that are definitions or definition-like are not stereotypes, one could thereby avoid accepting that the target proposition is a stereotype. (Similar cases include: French people are more likely than others to be born in France, or Irish passport holders are more likely than others to have been born on the Island of Ireland or have parents who were born there.)¹

A third potential response is to say that the example of the Durham teacher does not actually meet my definition of a stereotype because the association about where someone works is not a social attitude. Take the following definitions of social attitudes: “a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour” (Eagly and Chaiken 1993, 1). Or the definition of social attitude given by the *APA Dictionary of Psychology*:

1. A person’s general outlook on social issues and approach to his or her responsibilities
2. A person’s general disposition or manner towards other people (e.g. friendly or hostile)
3. An opinion shared by a social group (2023).

Social attitude, as I intend to use the terminology, is closest to 2 and 3. The attitude involves having a disposition towards people or an opinion of them, and it is a tendency to view people in a way that might favour or disfavour them by associating them with certain characteristics. It might be said that believing that someone is likely (or more likely than others) to

¹ If this option is taken, it will be important to think carefully about what is taken to be definition-like. For example, we would want to leave room for certain social attitudes that some people take to be definitional or definition-like to be stereotypes. For example, some people might think that women are by definition nurturing, but it does not seem that their viewing this attitude as definitional excludes it from being a stereotype. Consequently, it seems we should not like to say that what counts as a definition or definition-like is down to the beholder.

work in the North of England due to their working in Durham is not a disposition towards them, an opinion of them, or a tendency to view them in a way that may favour or disfavour them. Merely associating someone with working in a general geographical region based on knowledge that they teach in a more specific geographical region is not in and of itself having a disposition or an opinion towards them that might favour or disfavour them. This belief might lead to further opinions and positive or negative attitudes downstream, but the association itself does not represent an opinion or positive or negative attitude. Thus, it might be said that the association between teaching in Durham and working in the North of England is not a stereotype because it is not an expression of a disposition towards a person or people, an opinion of them, or a tendency to view them in a way that favours or disfavors them.

I am inclined to opt for a combination of the second and third of these responses: the association of Durham teachers with working in the North of England is not an expression of an opinion, a disposition towards a person or people, or a tendency to view a person or people in a way that will disfavour or favour them. Instead, it is a definition-like statement of a perceived evaluatively neutral fact. To be in the category of teachers who work in Durham one has, almost by definition, to work in the North of England. What consequence does this have for my definition of stereotypes? It may be altered as suggested above to stipulate that definitions and definition-like propositions are not stereotypes, but a large amount of the work of distinguishing cases of stereotyping and non-stereotyping comes from the stipulation that stereotypes are social attitudes.

2. Single/dual factor view

In HSDU I defend a multifactorial view of stereotyping, according to which there are multiple features of any act of stereotyping that determine whether the application of a stereotype is likely to increase or decrease the chance of an accurate judgement being made. In my defence of the multifactorial view, I reject two alternative views of stereotyping: the single and the dual factor view. According to the single factor view, a stereotype is likely to increase the chance of an accurate judgement being made when it is applied so long as it reflects an aspect of social reality. According to the dual factor view, the accuracy of the stereotype and the availability or lack thereof of unambiguous evidence determines whether the application of the stereotype increases or decreases the chance of an accurate judgement being made. My challenge to these views involves arguing that there are various other factors that determine whether an act

of stereotyping increases or decreases the chance of an accurate judgement being made. These factors include whether the stereotype is relevant in the context in which it is applied, and many other factors relating to how the person stereotyping accesses and processes case specific information (ambiguous or not) when engaging in stereotyping.

In this section, I am going to respond to some worries raised by Picinali and Arena in response to my critique of the single and dual factor views.

2.1 Critiquing the single factor view

When it comes to the single factor view, Picinali argues that this is so implausible as to be a straw man, and that no theorist would claim that stereotyping is likely to increase the chance of an accurate judgement being made. For example, it is a known fact that in racist societies characterised by racist police force Black people are overrepresented among those arrested for crimes like drug crimes, and this known fact means that people will be aware that a stereotype might reflect an aspect of social reality (i.e. arrest rates) but be unlikely to lead to a correct judgement. While I absolutely agree that there are facts like those gestured towards by Picinali that suggest that the single factor view is incorrect, it is nonetheless important to spell out a view like this and to challenge it. This is because it is likely that many people would defend specific stereotypes and acts of stereotyping precisely on the basis that there is a link between the stereotype and some aspect of social reality. When working on this type of issue, it is not enough to engage only with those who would develop well-formulated arguments with respect to stereotyping, it is important to engage with all those arguments that are likely powerful and widely endorsed. Picinali is right that the single factor view is weak and implausible, but to persuade me it is not worth considering it would be necessary to convince me that the argument has little or no power in society. Sadly, I suspect that this is untrue. Instead, my suspicion is that this type of position implicitly underlies much thought about stereotyping. As such, it is crucial to consider and reject it. In addition to this, the single factor view is used strategically as a jumping off point to develop the more plausible multifactorial view.

2.2 Critiquing the dual factor view

In discussion of the dual factor view, Picinali critiques my attempt to define what it is to be influenced by an accurate stereotype. I speak about a stereotype being accurate if it makes someone respond in a way that is fitting with accurate statistical information. For Picinali this reads as saying that the accuracy lies not in the generalisation but in the judgement

produced. Given that the dual factor view (and my general project at this point in HSDU) is to say when stereotypes are likely to lead to accurate and inaccurate judgements, a definition of an accurate stereotype as one that produces an accurate judgement would turn my claim that accurate judgements are more likely to be produced when accurate stereotypes are applied into a tautology.

To clarify, I do not mean to define accurate stereotypes as those that produce accurate judgements, but instead as those that reflect statistical information in a way that is accurate (for an understanding of what this might look like, see the work of Jussim and colleagues (2012) who argue that nearly all stereotypes reflect statistical information), leading people to make judgements that are reflective of that information. On my view a judgement could reflect background statistical information, via a stereotype, without being an accurate judgement. This is one of the main claims of the book: that even a person who makes a judgement that is in line with background statistical information (i.e. base rates) can nonetheless make errors, in particular, errors that are the result of their judgements being shaped by the background statistical information.

A second criticism that Picinali raises against the dual factor view relates to claims about relevance. One of my claims is that neither the single or dual factor views gives recognition to the ways that stereotypes can be, and seem likely to frequently be, applied when they are not relevant. Picinali criticises my work for not giving a definition of relevance. I was working with a commonsense understanding of relevance as meaning, roughly, pertaining to the truth of a judgement. Picinali argues that a common sense understanding of relevance does not seem to apply to at least one of my key examples of how stereotypes can be applied when not relevant:

A police officer approaches the car of a Black male, which has been pulled over for a minor traffic violation, e.g. one of his headlights is not working. The police officer asks the man to step out of the vehicle but he responds slowly and cautiously to the command. The police officer is offended at what he takes to be a threat to his authority. This triggers a stereotype associating the innocent man with crime; the police officer evaluates the man as a criminal and treats him with hostility; and this leads to an escalation of tension and hostility between the two individuals. The stereotype associating Black people with crime is triggered although the Black man has not committed a crime, only a minor traffic violation (HSDU, 46)

Picinali argues that the stereotype could be relevant in this case because after stopping the car the police officer may become suspicious that the driver is or will commit a crime. Then the stereotype about social group and criminality would be relevant. I put aside the latter point about whether the stereotype becomes relevant if a suspicion arises about there being a high chance that the driver would engage in criminality. I have serious doubts about this: why would the stereotype be relevant when the police officer has access to case specific information about the driver? I do not need to develop this doubt any further, however, because there is another problem with this critique: it requires adding detail to the example and thereby changing it. The example is designed to show that there are likely to be many cases in which a stereotype is triggered when not relevant. This point is backed up by psychological findings speaking to the question of when stereotypes tend to be triggered (i.e. when there is a challenge to the status quo or someone has a wounded ego) and the example, as presented. In the example, a Black person is stopped for a traffic violation. The stereotype associating Black people more strongly than white people with criminality does not pertain to the truth of the judgement of the driver's situation. It is not relevant. We can try to construct similar cases, adding detail to show how in different cases stereotypes might be relevant. However, this does not undermine the point that there are conditions, seemingly many, in which a stereotype is likely to be triggered but not relevant on any commonsense understanding of the term.

Arena raises a similar, but importantly distinct, worry about the idea that relevance is a factor that determines if stereotyping helps or hinders judgement that is not captured by the single or dual factor view. Arena provides three potential definitions of relevance:

- (1) Relevance as meaning that the individual is within the scope of the stereotype.
- (2) Relevance as “non-spurious statistical correlation”.
- (3) Relevance as “argumentative force”.

Arena suggests that the dual factor view captures (1) and (2). So if by focusing on relevance I meant to show that stereotypes can be applied to people outside the scope of the stereotype, or when there is at best a spurious statistical correlation between group membership and a target characteristic then I would not be presenting a challenge to the dual factor view. In fact, I did not mean to refer to (1) or (2), but instead, as mentioned above, something like “pertains to the truth of”. My point is that a person may fall within the scope of the stereotype, in the sense that they have the social identity that the stereotype makes an association with, and the stereotype can still lead the person stereotyping astray independently of

whether there is a non-spurious correlation between being a group member and having the characteristic. Even if there were a non-spurious correlation between being member of group G and having trait T, and person x falls into category G, the stereotype associating G with T may be applied to x under conditions where the stereotype is uninformative, that is, the stereotype does not pertain to the truth of whether G is displaying characteristic T. What I am getting to is closest to (3): the idea that the stereotype does not have force in the context in which it is applied. But it is not argumentative force, as in persuasive force, but instead something more like evidential force, whether evidence pertains to the truth of the matter at hand.

One final point on the dual factor view. I argue that there are various ways that stereotypes can deceive us by leading to a distorted response to case specific information. In the terminology I use later in the book: there are various ways that stereotypes dispose us to respond poorly to information that is available in our environment. Arena suggests that the dual factor view can accommodate this thought:

[G]iven the emphasis that the Dual factor view puts on case-specific information, (...) the cases of stereotyping pointed out by the examples would be considered by the Dual factor view as lacking epistemic quality, given that the ambiguity of individual information was not adequately ruled out. (Arena this issue, 18)

While I appreciate that the effects that I describe are ways that the information about an individual is not adequately assessed when stereotyping occurs, my aim in presenting the multifactorial view as an alternative to the dual factor view is to emphasise that even if one's external environment provides high quality information that information might be hard to access and properly process precisely because one harbours a stereotype.

3. Epistemic benefits of egalitarian beliefs

One of the main goals of the book is to outline how stereotypes can lead us to misperceive and misunderstand people and events involving people. As such, via the multifactorial view, the book outlines various epistemic pitfalls that are associated with stereotyping. Building on this picture of the various epistemic pitfalls that are associated with harbouring stereotypes, I argue that having false egalitarian beliefs can bring overall epistemic benefits, outweighing the epistemic costs of their falseness. The egalitarian

beliefs bring epistemic benefits because they prevent people from being subject to the pitfalls associated with stereotyping. Picinali and Arena both object to the idea that egalitarian beliefs guarantee epistemic benefits, and the avoidance of the epistemic costs. Picinali gives the example of someone who endorses an attitude according to which both men and women are unlikely to have scientific experience. The attitude may lead the person to make epistemic errors, e.g., misinterpreting ambiguous information, misremembering scientific achievements, etc. For Arena,

The point here is that the distortion of case-specific information is a consequence of basing the assessment of individual traits on a generalisation, regardless of whether it has an egalitarian or non-egalitarian content. (Arena this issue, 20)

These are interesting and important points. In reply, I would first re-iterate the strength of the claim that I defend. It is that *often*, and likely *more often than not*, the epistemic benefits of a false egalitarian attitude will outweigh the epistemic costs of the falsity of the attitude. For example, for some people it may be better from an epistemic perspective to falsely believe that men and women are equally likely to have scientific expertise. I think that this is plausible because stereotyping brings the risk of a raft of epistemic errors, including memory errors, misinterpreting ambiguous evidence, assuming similarities when they do not exist, failing to notice similarities when they do, failing to give adequate credibility to testimony, etc. Having an egalitarian attitude can reduce the risk of at least some of these errors, thereby increasing the chance of true beliefs, understanding, and so forth. What I do not claim is that egalitarian attitudes guarantee that every one of the epistemic errors will be avoided, or that the egalitarian attitudes cannot bring similar errors. So, it is compatible with my view that egalitarian attitudes can bring epistemic costs and that the egalitarian attitudes do not guarantee the avoidance of epistemic error.

Nonetheless, it would be remiss not to acknowledge how Picinali and Arena's comments have challenged me to think further about how egalitarian attitudes can bring epistemic costs. It certainly seems right that if you are a committed egalitarian, you may be prone to interpreting ambiguous evidence in a way that is consistent with your egalitarian beliefs, for example.

However, when considering other epistemic pitfalls associated with stereotyping, it seems less plausible that egalitarian beliefs will have the same negative epistemic effects as stereotypes. This is because the negative effects of the stereotyping are specifically associated with the process of categorisation of individuals into different social groups. An

egalitarian judgement that does not encourage seeing members of the groups differently seems less likely to have the same effects. Take, for example, the way that classifying someone as a member of a minority group can lead them to be viewed as more similar than they really are to other members of the same group (Bartsch and Judd 1993; Hewstone, Crisp, and Turner 2011). Or take the way that those classifying individuals as a member of one group (e.g. women scientists) may fail to notice similarities between those individuals and members of other groups (e.g. men scientists) (e.g. Tajfel 1981). It seems unlikely that there will be similar effects that occur because of egalitarian beliefs that emphasise similarities, or include claims that apply to all relevant groups (e.g. both men and women who are scientists).

Something similar can be said about the relationship between stereotyping and memory (cf. Arena this issue). As I discuss in my HSDU, psychological studies suggest that people often remember information consistent with a stereotype better than other information, and this effect is explained in terms of the role of social schemas or expectancies (e.g. Rothbart et al. 1979; Fiske and Linville 1980). It is argued that information that is consistent with a social schema is often more easily stored and retrieved from memory. Where there is a stereotype relating women to being less likely to have scientific expertise than men this will be a part of the social expectancies or social schema “WOMAN”. This social schema is likely to be activated in response to women scientists, and to shape the way that information about them is processed and stored. It is far from clear, in contrast, that if someone has the egalitarian attitude that both men and women are likely to lack scientific expertise, that lacking scientific expertise will be a part of the social schema for either men or women. Or that the social schema of “MAN” or “WOMAN” will be activated in response to any individual scientist, leading to schema-consistent memory effects. What seems more likely to happen is that all scientists will be viewed as exceptions to the general rule that men and women are unlikely to be good at science. For one final example, in HSDU, I discuss Kristie Dotson’s (2011) work on testimonial smothering, that is, on the way that people may choose to suppress risky testimony, especially if they fear that it will sustain or compound existing stereotypes that they take others to possess. There seems to be good reason to think that a woman is likely to suppress risky testimony about a scientific matter to prevent themselves from compounding the stereotype that women are less likely to have scientific expertise than men. In contrast, there seems to be significantly less risk of someone choosing not to speak about a scientific matter for fear of compounding the stereotype that men and women are both unlikely to have scientific expertise.

In sum, then, egalitarian attitudes do not appear to bring the same risk of epistemic pitfalls as stereotypes because many of the pitfalls of stereotyping seem to be closely tied to features of social categorisation. Because there are some epistemic pitfalls more strongly associated with stereotyping than egalitarian beliefs, false egalitarian beliefs are likely to often bring more epistemic benefits than stereotypes that they would replace. This is all that is needed to support my claim that *often*, and likely *more often than not*, having a false egalitarian belief can be best from an epistemic perspective.

4. Beyond stereotyping beliefs

Although I maintain that false egalitarian beliefs can be better from an epistemic perspective because they can avoid significant epistemic costs, the idea that egalitarian beliefs can also dispose people to fail to respond appropriately to information that they encounter downstream brings me to my next point.

In HSDU I present a new approach to evaluating stereotypes and stereotyping along the epistemic dimension: evaluative dispositionalism. Evaluative dispositionalism encourages people to focus on the epistemic dispositions, that is, dispositions to respond one way or another to evidence, that a person has displayed when acquiring a stereotype *and* those epistemic dispositions that a person has due to harbouring the stereotype. But, as pointed out by Picinali, many of the examples that I use in HSDU to illustrate the nature of an epistemic disposition are unrelated to stereotyping (e.g. I mention how a person's beliefs about a football team can shape their responses to evidence about their performances). Picinali's observation points towards a broader ambition that I have: to apply evaluative dispositionalism to other, non-stereotyping, beliefs. For example, I think that the evaluative dispositionalist approach could be fruitfully applied to the core beliefs held by members of echo chambers; that it would be valuable to evaluate both the epistemic dispositions that people have displayed in entering the echo chamber and the way that they are disposed to respond to evidence once they have the core echo chamber beliefs. The evaluative dispositionalist approach could also be applied to egalitarian beliefs, where those dispose people to respond poorly to evidence.

Considering the issues raised by the commentators on my book in this symposium (see sections 5 and 6 below), it would be interesting to also consider whether we should move away from evaluating the epistemic standing of individual beliefs, towards considering how individual beliefs

interact with other beliefs, situational factors, and personality types to dispose people to respond well or poorly to evidence.

5. Which disposition?

In my defence of evaluative dispositionalism in HSDU I suggest that it would be useful advice to “check your dispositions” in relationship to stereotyping—rather than merely considering how a stereotype is formed, one ought to consider how one is disposed, due to possessing the stereotype, to respond to evidence about individual people or events. This is because people can come to harbour stereotypes in better or worse ways—displaying poor dispositions or merely experiencing the misfortune of being in a hostile environment—but people can also be disposed to respond in better or worse ways to relevant information once they harbour the stereotype.

Saul suggests that it will be difficult to apportion praise or blame to people on this type of approach because we cannot possibly know what dispositions a person has. This problem is especially acute given how many dispositions people have in relation to any specific belief. Saul calls this *the very many dispositions argument*. The very many dispositions argument highlights a serious issue. It will not always, or perhaps often, be possible to be sure about all dispositions that someone has because of believing any proposition (or harbouring any implicit stereotype non-propositionally). It will be difficult in any specific case to identify how a person is disposed to respond due to the stereotypes that they harbour, especially to the level of certainty that might be required for justified praise or blame. Nonetheless, I believe that there are good reasons for thinking that it is still valuable to focus on dispositions.

First, if you are aware of your own tendency to stereotype, or that you are likely to stereotype, you can make efforts to reflect on whether stereotyping is likely to be leading you astray in *any specific context*. It is not necessary to be able to identify all dispositions that might be manifest in any context to do this, so the very many dispositions argument does not come into force. Notwithstanding the shortcomings in people’s awareness of what stereotypes they harbour and when stereotypes are likely to be triggered, people who are informed by relevant psychological findings, theoretical accounts of stereotyping, personal accounts of being on the receiving end of stereotyping, etc.—the types of information discussed in HSDU—can reflect on the likely effect of stereotypes on the dispositions that they are may display in a particular context. For example, they can reflect on why they noticed certain information, whether there was other

information they did not notice, whether there are gaps in their memories, and so forth. My view encourages and supports people to engage in this type of reflection on the dispositions that they might display by highlighting the types of dispositions that stereotypes bring.

Second, a focus on the dispositions that *other people* are likely to display in any specific context due to stereotyping can lead to appropriate skepticism towards the beliefs that they articulate relating to groups that they are likely to stereotype. One does not have to know for certain which dispositions a person holds to factor in that they might be misled by stereotypes, and that it is therefore worth seeking out evidence that could confirm or disconfirm whether they are being misled. It can be useful to be aware of the types of dispositions that a person might have due to stereotyping, to probe them to test for likely effects. For instance, let us say that someone claims that a member of a minority group in your workplace has not been contributing as they should. One might seek out information about what specifically that person has noticed and remembers in relation to the minority group member, the context in which they have interacted, and in which their memories about the minority group member were formed. This information can provide guidance about whether the person is likely to be displaying a disposition not to properly encode, process and recollect information about the member of the minority group, due to stereotyping.

The argument here has similarities to one that Saul (2013) provided when defending what has become known as “Saulish skepticism” (Antony 2016) in relation to implicit bias. Saul argues that evidence of implicit bias provides reason to adopt a skeptical attitude towards beliefs about social actors and objects because those judgements could easily, unbeknownst to the person judging, be influenced by irrelevant and distorting implicit stereotyping. I, like many others (e.g. Fricker 2007; Antony 2016), emphasise that stereotypes can operate, either explicitly or implicitly, to supply information that is relevant to a judgement. However, I also argue that stereotypes—even those that reflect reality—can dispose us to respond poorly to evidence. This suggests that we ought to adopt a skeptical stance towards beliefs that may have been shaped by stereotypes, carefully considering how stereotypes might have operated to influence how the beliefs have been formed by shaping the dispositions of the believer.

Third, while it will be difficult to discern for any individual which dispositions they possess due to stereotyping, there is an important lesson to be learnt from this problem. The lesson is this: it is important to be aware that information is not equally safe in anybody’s hands. Certain information about social groups can be useful and informative in the hands

of some people while damaging in the hands of others. The main difference-maker could be the dispositions that those people possess due to harbouring the stereotype. Not only this, it may be difficult or impossible to know all of the dispositions that a person could harbour due to stereotyping. It is therefore important to be cautious about whom to trust in relation to specific information—for example, information about a person's mental health condition—because one cannot be sure how the information may lead any specific person to be disposed to respond to evidence that they might encounter.

6. More radical implications of evaluative dispositionalism

As mentioned in sections 4 and 5, according to the view proposed in chapter 8 of HSDU, evaluative dispositionalism, one ought to evaluate a stereotyping belief by considering both the dispositions displayed forming the belief and the dispositions possessed due to harbouring the belief. This pluralistic approach involves considering both the causal history of the stereotyping belief (the dispositions displayed when coming to possess a belief), and the consequences of believing (the dispositions held as a result).

Saul does not object to the proposal that it is worthwhile considering the causal history and consequences of stereotyping, but suggests that my argument implies something more radical than evaluative dispositionalism, which focused on the causal history and consequences *of a particular stereotyping belief*. For Saul my arguments suggest a more intriguing possibility: that we should not be epistemically evaluating any single belief on its own.

Saul gives the example of two individuals, Betty and Caleb, both of whom take a training session on bias. While Caleb is attentive, takes a handout, reads it carefully, and attempts to prevent a specific stereotyping belief from influencing his cognition downstream, Betty does not even look the handout. For Saul, it is not any stereotyping belief alone that is an apt object of epistemic appraisal. Other beliefs, such as the belief about whether it is worthwhile looking at the handout, are also apt, and perhaps even more apt, for praise or criticism. Saul also gives the example of Edith and Dorinda, who both believe that if a man and woman seem equally qualified on paper the woman is in fact better qualified because she will have achieved what she has despite facing barriers that the man has not faced. Although Edith and Dorinda are similar in this way, Edith but not Dorinda understands the barriers faced by black men. This leads Edith but not Dorinda to factor in a man's race when considering the barriers faced

by the two individuals. Saul takes both these examples to show that “what dispositions people have arising from any belief depends on many other facts about them—including, crucially, what other things they believe” (Saul this issue, 64), and, furthermore, to suggest that when it comes to stereotyping it is not just a specific stereotyping belief that ought to be evaluated.

A similar thought emerges from Picinali’s contribution. Picinali takes one of the examples from the book, of Nora and Ned, to show that the focus on any specific stereotyping belief seems to narrow. In the example, both Nora and Ned harbour the stereotyping belief that men are more likely than women to have scientific expertise. However, only Ned makes a catalogue of errors due to the stereotype: e.g. memory errors, misinterpreting ambiguous information, and so forth. Because both individuals harbour the stereotype, Picinali argues, the stereotype itself can only have a limited causal role in determining the dispositions, otherwise both Nora and Ned would display the same dispositions. The causal story is more complex than evaluative dispositionalism suggests.

These comments have been extremely helpful in clarifying my thinking about stereotypes and dispositions. One of the main motivations behind the HSDU project was to show that a simplistic picture of stereotyping cannot work because of the variety of ways that believing a stereotype can shape our cognition as well as our action. There will be individual differences in the way that stereotypes shape responses to evidence. Some of these differences are due to situational factors, e.g. time, cognitive load, amount of information to process, others relate to personality, and others further to the wider set of beliefs that are held. Given this, it seems right, as Saul and Picinali suggest, that stereotyping beliefs should not be considered and evaluated in isolation. An epistemic evaluation of an act of stereotyping should consider how the stereotyping belief is likely to lead a specific person to be disposed to respond to information given other facts about them, including the other beliefs that they hold. This is not to diminish evaluative dispositionalism as an approach to stereotyping. But it is to acknowledge that it may be more important than HSDU has emphasised to consider other factors that have a causal role in determining how stereotypes deceive us, and what they mean for how we might be disposed to respond to evidence. As Saul suggests, my argument for evaluative dispositionalism provides some support for the radical idea that when epistemically evaluating people’s stereotyping, we ought not to only focus on any specific stereotyping belief, but a far wider range of phenomena.

7. Stereotypes, reality and testimonial injustice

Leonie Smith's contribution encourages us to delve deeper into the relationship between stereotypes that appear to reflect reality and testimonial injustice. The terminology of testimonial injustice was introduced by Miranda Fricker (2007) to capture cases where identity prejudice leads the testimony of members of marginalised groups to be systematically given reduced credibility. Fricker describes prejudice as involving an epistemically culpable failure and affective investment. In HSDU, I describe cases where people suffer harm because they are given less credibility than they deserve due to stereotypes relating to their social identity, but where the stereotyping does not seem to meet Fricker's definition of prejudice because it does not involve affectively invested epistemic culpability: the stereotype is acquired in response to the information available in one's environment. I often focus on the stereotype associating scientific expertise more strongly with men than women. A person could acquire this stereotype by being well-informed about levels of training across gender groups rather than due to any affective investment in the stereotype. In commentary on these ideas, however, Smith encourages us to think further about the relationship between stereotypes like the science stereotype and testimonial injustice. Smith suggests that, properly understood, prejudice can be viewed as having an important role in the types of case I focus on.

Smith describes several ways that these types of case could involve prejudice and/or epistemic culpability, and so be classified as testimonial injustices.

- (1) The cases may be said to involve prejudice because prejudice shapes the social backdrop, e.g. levels of expertise among men and women, that makes the stereotype accurate.
- (2) The people in the cases may have internalised prejudiced beliefs due to the prejudiced social backdrop, so only appear not to be prejudiced.
- (3) Prejudice could be present in the inappropriate application of stereotypes in contexts where evidence suggests that they do not apply.
- (4) The cases could involve prejudice because although there appears to be an accurate and well-supported stereotype in operation there is in fact an inaccurate one that is poorly supported by the evidence.

This analysis provides a more fine-grained taxonomy of cases than is provided in HSDU. It presents several distinctive ways of conceptualising

prejudice and its role in credibility assessments. It thereby points towards different ways that the condition that testimonial injustice involves prejudice could be met. However, I would add the following observations to Smith's suggestions.

First, while it is important to recognise that prejudice can form the backdrop to people having stereotypes that are accurate, there is still value in distinguishing between cases where people dismiss other people as lacking credibility due to being prejudiced, and those where people dismiss others as lacking credibility because they grew up in societies shaped by prejudice. One way to make this distinction is to stipulate that there is only testimonial injustice when prejudice directly causes the credibility deficit and not where prejudice only indirectly influences the credibility judgement by shaping society so certain stereotypes are accurate. If, on the contrary, we were to say that there is testimonial injustice in all cases where there is a backdrop of prejudice then this important distinction will be lost.

Second, while it is highly plausible that many people in prejudiced societies internalise prejudice, this does not mean that there are not also people who engage in stereotyping, and subsequently harm others by giving them less credibility, without having internalised prejudice. As discussed in HSDU, psychological evidence strongly suggests that stereotypes impact how people respond to evidence, including about people's credibility, regardless of whether those stereotyping endorse prejudicial beliefs or have any affective investment in them. It may be that these responses to evidence are indirectly due to prejudice in society, reflecting, for example, prejudicial social structures, but this does not equate to them being due to internalised prejudice. In HSDU I encourage readers to recognise the epistemic costs that can follow from this type of stereotyping, absent the direct role of prejudice. I would therefore stress the importance of not conflating the claim that some people are likely to internalise prejudice in prejudicially structured societies with the claim that all harmful credibility deficits are due to internalised prejudice.

Third, it is important to distinguish two claims: (i) prejudiced attitudes can be *reflected in* the way that people apply a stereotype, i.e. whether they apply it out of context, (ii) prejudice can be *constituted by* people applying a stereotype out of context. (i) certainly seems to be true. Prejudice can lead people to apply generalisations more broadly than they should, for example, applying the stereotype that scientific expertise is more common among men than women when it should not be applied, to a trained woman scientist. But this observation only establishes that some cases in which people apply stereotypes out of context, or when they are irrelevant, are ones where prejudice has been operational. It does not establish that in all

cases where people apply stereotypes where they are irrelevant this is due to prejudice. It therefore leaves open the possibility that some cases in which people apply a stereotype when it is irrelevant occur in the absence of prejudice having a direct role.

On the other hand, it might be argued that when people apply stereotypes out of context this *constitutes* prejudice. This would suggest that wherever stereotypes are applied out of context, like they are to the woman scientist, there is prejudice. So at least some of the examples that I use to support the claim that there can be credibility deficits in the absence of prejudice having a direct role—i.e. those where stereotypes lead to credibility deficit only because they are applied out of context—could not be used in this way. They would be cases where prejudice has a direct role. This would weaken my case in support of the claim that there are credibility deficits due to stereotyping that do not constitute testimonial injustice because there is no prejudice involved. However, making this move would also involve significantly revising any conception of prejudice that does not define it in terms of stereotypes being applied out of context, which I take would include many commonsense conceptions.

Finally, I agree there will be cases where judgements appear to be underpinned by accurate stereotypes, but inaccurate stereotypes are in fact operational. One general challenge associated with evaluating someone's acts of stereotyping is to pin down the content of the generalisation that they are applying. However, I take it that rather than counting against my position in HSDU, these observations provide additional reason to take evaluative dispositionalism seriously. Evaluative dispositionalism provides a framework that can be applied to evaluate acts of stereotyping even when it is difficult to establish beyond doubt the content of the stereotype, or whether the content accurately reflects reality. You can seek evidence about whether the person engaging in stereotyping seems to be displaying the dispositions that would be associated with stereotyping of the type that is suspected.

8. Normative stereotypes

Arena points towards an important distinction that is not covered in my book, but which is of importance to discussion of stereotyping. This is the distinction between descriptive and normative stereotypes:

In the case of *Price Waterhouse v. Hopkins*, the American Psychological Association's *amicus curiae* noted the importance of distinguishing between descriptive and

normative stereotypes about women. The authors of the *amicus curiae* claimed that: “descriptive stereotypes characterize women in a way that undermines their competences and effectiveness; normative stereotypes label women whose behaviour is inappropriately masculine as deviant”. (Arena this issue, 21)

Normative stereotypes are stereotypes that dictate how people who are categorized under them ought to behave, characterising ways that it is deemed appropriate for them to behave. For example, a normative stereotype might dictate that it is appropriate for women to be co-operative. Arena points out how normative stereotypes attempt to shape behaviour and can shape the social world to fit the stereotypes. Normative stereotypes, for Arena, can have dangerous epistemic effects, as people may abandon their epistemic goals, focusing on disapproving of people who do not conform, and “constructing the facts in such a way as to make it possible to inflict some type of punishment” (*ibid.*, 21), e.g. denying that a victim is a victim on the basis that she did not conform to stereotypical behaviour. On a similar note, Wade Munroe (2016) has argued that there can be “prescriptive credibility deficits” that occur when a speaker who fails to meet a relevant prescriptive stereotype (e.g. a stereotypically feminine speaking style) are assigned a lower level of credibility than they would be otherwise. The prescriptive aspect of stereotyping is not an aspect that I explore in the book, but I agree that it is an important phenomenon.

9. Moral encroachment

In a brief section in chapter 5 of HSDU, I compare my argument to moral encroachment views. The basic claim of moral encroachment views is that moral factors can determine what one is justified in believing, can believe rationally, or knows. Some moral encroachment theorists argue that the moral encroachment view can dissolve an epistemic-ethical dilemma that has been argued to be posed by stereotypes and stereotyping (Basu 2019a, 2020; Basu and Schroeder 2018). The purported epistemic-ethical dilemma is that where stereotypes reflect aspects of social reality—e.g. base-rate information about rates of arrest across different social groups—it can be epistemically beneficial to apply these stereotypes, but at the same time unethical (Gendler 2011; Mugg 2013). One faces a dilemma between achieving one’s epistemic goals or one’s ethical goals. Some moral encroachment theorists have argued that a dilemma like this does not emerge when it comes to (at least many) stereotypes: the stakes involved in situations in which people might stereotype raise the evidentiary standards, so a high level of evidence is required to be justified in

believing, to be rational, or to know. However, those engaging in stereotyping will fail to meet these standards. They will therefore not be justified because the moral stakes of the situation have raised the evidentiary standards beyond those attained by the believer. The correct thing from the epistemic perspective will be the ethical thing, i.e. not stereotyping.

In HSDU, I present an alternative approach to the epistemic-ethical dilemma. I argue that the situation that people face in relation to stereotyping is far more complex than the simplistic description of the epistemic-ethical dilemma suggests. Sometimes applying stereotypes that reflect reality can be epistemically costly rather than beneficial, because of the epistemic pitfalls associated with stereotyping. Sometimes it can be ethically required, for example, if the stereotype associates members of a particular group more strongly than others with certain medical or social conditions that they require help with. Sometimes epistemic and ethical goals conflict, but sometimes they concur, with both epistemic and ethical goals being achieved either by stereotyping or not doing so. I argue that this analysis of the complex interplay of epistemic and ethical demands of stereotyping provides things that are missed by moral encroachment views that focus narrowly on the idea that some stereotyping can never meet the high evidentiary standards set by the stakes of social situations in which stereotyping occurs.

In Picinali's contribution to the symposium he focuses on one specific idea relating to this discussion of moral encroachment. He argues that moral encroachment views can explain how some stereotyping may be epistemically acceptable, taking this to undermine the claim that my approach is preferable. Picinali applies a formal framework, initially proposed to model pragmatic encroachment views, to argue that

[T]here will be situations in which the false negative has such a high moral cost that the threshold for outright belief (and, hence, for knowledge) will be relatively low; sufficiently low to be satisfied by stereotyping. In the medical example in which a false negative (missed diagnosis) leads to a quick death and a false positive (false diagnosis) leads to some health benefits and mild side effects, the doctor may well be morally warranted to follow a stereotype linking members of a group to which the patient belongs with the medical condition at issue (more strongly than non-members). (Picinali this issue, 51-52)

This is not the place to dig into the details of the framework Picinali proposes. It may be possible to develop a formal model that assigns a value

to different outcomes (e.g. correct diagnosis, incorrect diagnosis), reflecting the morality of the outcomes, and sets a probability threshold that determines when it is justified to act as if a proposition, i.e. a stereotype, is true. It may be that at times the values are set such that one could be said to be justified in acting as if the stereotype is true, given, for instance, the strong moral demand to achieve an outcome that this would facilitate. However, the development of this model would only get us so far in understanding the costs and benefits of stereotyping, and how we can do our best in relation to stereotyping. What my approach suggests is that acting as if a stereotype is true can at the same time bring benefits and very significant costs. Adopting a fine-grained approach to understanding stereotyping, like the one proposed in HSDU, allows us to focus on both the costs and the benefits, recognising this complexity. Ideally, it would enable people to reflect on their practices, to harness the benefits while avoiding at least some of the costs of stereotyping. A formal model that simply delivers a result that either people are or are not warranted to act as if a stereotype is true would obscure the complexity of the situation, which, I argue, needs to be faced head on.

Acknowledgments

I'm very grateful to Marina Trakas for organising this symposium and to all the contributors for their thoughtful comments on my book.

REFERENCES

- Antony, Louise M. 2016. "Bias: Friend or Foe? Reflections on Saulish Skepticism." In *Implicit bias and philosophy. Volume 1: Metaphysics and Epistemology*, edited by Michael Brownstein and Jennifer Saul, 157-190. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198713241.003.0007>.
- APA Dictionary of Psychology. 2023. "Social Attitude." URL: APA Dictionary of Psychology.
- Bartlett, Frederic Charles. 1995. *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.
- Bartsch, Robert A., and Charles M. Judd. 1993. "Majority—Minority Status and Perceived Ingroup Variability Revisited." *European Journal of Social Psychology* 23 (5): 471-483. <https://doi.org/10.1002/ejsp.2420230505>.
- Basu, Rima. 2019. "The Wrongs of Racist Beliefs." *Philosophical Studies* 176 (9): 2497-2515. <https://doi.org/10.1007/s11098-018-1137-0>.

- Basu, Rima 2020. "The Specter of Normative Conflict: Does Fairness Require Inaccuracy?" In *An Introduction to Implicit Bias: Knowledge, Justice, and the Social Mind*, edited by Erin Beeghly and Alex Madva, 191-210. New York, NY, USA: Routledge.
- Basu, Rima, and Mark Schroeder. 2018. "Doxastic Wronging." In *Pragmatic Encroachment in Epistemology*, edited by Brian Kim and Matthew McGrath, 181-205. New York, NY, USA: Routledge.
- Beeghly, Erin. 2015. "What is a Stereotype? What is Stereotyping?" *Hypatia*, 30 (4): 675-691.
<https://doi.org/10.1111/hypa.12170>.
- Blum, Lawrence. 2004. "Stereotypes and Stereotyping: A Moral Analysis." *Philosophical papers* 33 (3): 251-289.
<https://doi.org/10.1080/05568640409485143>.
- Dotson, Kristie. 2011. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26 (2): 236-257.
<https://doi.org/10.1111/j.1527-2001.2011.01177.x>.
- Eagly, Alice H., and Shelly Chaiken. 1993. *The Psychology of Attitudes*. Harcourt Brace Jovanovich College Publishers.
- Fiske, Susan T., and Patricia W. Linville. 1980. "What does the Schema Concept Buy Us?" *Personality and Social Psychology Bulletin* 6 (4): 543-557. <https://doi.org/10.1177/0146167280640>.
- Miranda Fricker. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Gendler, Tamar Szabó. 2011. "On the epistemic costs of implicit bias." *Philosophical Studies* 156: 33-63.
<https://doi.org/10.1007/s11098-011-9801-7>.
- Hewstone, Miles, Richard J. Crisp, and Rhiannon N. Turner. 2011. "Perceptions of Gender Group Variability in Majority and Minority Contexts." *Social Psychology* 42 (2): 135-143.
<https://psycnet.apa.org/doi/10.1027/1864-9335/a000056>.
- Jussim, Lee. 2012. *Social Perception and Social Reality: Why Accuracy Dominates Bias and Self-fulfilling Prophecy*. New York: Oxford University Press.
- Kahneman, Daniel. 2011. "Thinking, Fast and Slow." *Farrar, Straus and Giroux*.
- Mugg, Joshua. 2020. "How not to Deal with the Tragic Dilemma." *Social Epistemology* 34 (3): 253-264.
<https://doi.org/10.1080/02691728.2019.1705935>
- Munroe, Wade. 2016. "Testimonial Injustice and Prescriptive Credibility Deficits." *Canadian Journal of Philosophy* 46 (6): 924-47.
<https://doi.org/10.1080/00455091.2016.1206791>
- Rothbart, Myron, Mark Evans, and Solomon Fulero. 1979. "Recall for Confirming Events: Memory Processes and the Maintenance of

- Social Stereotypes.” *Journal of Experimental Social Psychology* 15 (4): 343-355. [https://doi.org/10.1016/0022-1031\(79\)90043-X](https://doi.org/10.1016/0022-1031(79)90043-X).
- Saul, Jennifer. 2013. “Scepticism and Implicit Bias.” *Disputatio* 5 (37): 243-263. <https://doi.org/10.2478/disp-2013-0019>
- Tajfel, Henri. 1981. *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge: Cambridge University Press.