

Epistemic injustice and implicit bias

Jules Holroyd & Kathy Puddifoot

Abstract:

What is the relationship between knowledge and power? Philosophers have described various kinds of *epistemic injustice* – ways in which our practices for seeking and sharing knowledge can be unjust. Who is treated as trustworthy? Who receives credit for ideas? Whose hard work in advancing knowledge is recognized? Such judgements may be affected by power, and influenced by stereotypes. A further set of questions concerns: what concepts or interpretive resources do we have for making sense of the world? Whose experiences are better or worse “fits” for these concepts? Who gets to shape these concepts? Again, the answers to these questions may reflect social dynamics related to power and stereotypes. Reflections on the film *Hidden Figures* illustrates these claims. We tease out how implicit biases might be implicated in these injustices. This prompts reflection on the kinds of social and institutional changes needed to address epistemic injustices.

How, if at all, do knowledge and social power relate to each other? A common sense thought is that our practices of inquiry and knowledge seeking have little to do with politics and social power: we simply find facts and build knowledge. But feminist philosophers have long argued that power and politics affects our knowledge-seeking practices. What is known and how inquiry proceeds are thoroughly inflected by social dynamics. Moreover, a strand of epistemology called social epistemology has emphasized that human practices of seeking knowledge often have a social dimension. We depend upon each other as sources of knowledge and understanding, in order to access resources for developing our knowledge, and in order that we can contribute to shared understandings of our world. This is to say that our knowledge seeking practices – our epistemic practices – are social. This helps us to see how social power can affect knowledge-seeking: because our epistemic practices are social, the kinds of power dynamics that we find in social relations can impact on them. In particular, unequal power relations can impact our knowledge seeking endeavors.

In this chapter, we set out some of the ways that social power can affect knowledge, identifying instances of what philosophers have called ‘epistemic injustice’. This notion is explored by Patricia Hill Collins in *Black Feminist Thought* who articulated the ways that ‘power relations shape who is believed and why’ (2000, 270). Collins articulates how important knowledge is therefore marginalized. More recently, epistemic injustice has been characterized as ‘consisting, most fundamentally, in a wrong done to someone specifically in their capacity as a knower’ (Fricker, 2007, 1). The ability to produce knowledge and contribute to inquiry is a fundamental part of what it is to be human, Fricker argues. As we will see, there are various ways in which one might be harmed in this capacity. Here we focus on how racism and sexism can contribute to epistemic injustice.

The film *Hidden Figures* provides a focal point for our analysis of epistemic injustices. We illustrate some key forms of epistemic injustice with reference to the experiences described of some of the black women — Katherine Goble (now Katherine Coleman Johnson Goble), Dorothy Vaughan and Mary Jackson—working at NASA (National Aeronautics and Space Administration) in the United States in the early nineteen sixties. These women made enormous contributions to the mathematical projects required to send the astronaut John

Glenn into space, as described in the book by Margot Shetterley, and depicted in the 2016 film of the same title, *Hidden Figures*. These contributions were made notwithstanding that, given the intersecting oppressions they faced, they each experienced various kinds of injustices, including epistemic injustice, as we describe here. Whilst an important retelling of their stories, we note the ways that the film itself may also be implicated in certain epistemic injustices.

As we describe some of the different ways in which epistemic injustice may be perpetrated, we present the case for supposing that implicit biases might contribute to these forms of epistemic injustices. Since implicit biases cannot be overcome simply by intending to avoid bias, this poses difficult questions about how to address epistemic injustices, and in particular, what sorts of collective and structural strategies need to be employed to confront them (see Davidson & McHugh, this volume; Ayala-Lopez, this volume; Madva, this volume).

Kinds of epistemic injustice

There are various ways in which we participate in social practices as knowers and knowledge seekers. We get knowledge from others and share it with them; we try to access resources in order to develop our understanding of the world further; and we try to make and be recognized for, contributions to our shared understanding.

There are some obvious ways in which societies that are structured unequally will lead to epistemic injustices and hinder knowledge production. One protagonist of *Hidden Figures*, Mary Jackson, showed great flair for engineering, but could not work as an engineer at NASA without formal qualifications. These qualifications were only achievable by attending classes at a white-only school. Racial segregation posed a barrier to Jackson – and many others – insofar as her access to knowledge, and the markers of knowledge (the qualification) were restricted due to race.

Jackson's determination led her to court to petition for access to the necessary classes. Even when she gained access to the necessary classes, she did so after expending energy and confronting indignities that white students did not have to face. This is a clear case, then, in which racism hinders access to knowledge. But philosophers have identified other, perhaps less obvious, ways in which epistemic injustices might manifest. We outline these below, showing how the dynamics of racism and sexism of 60s America led to various kinds of epistemic injustice against the striving mathematical engineers Goble, Vaughan and Jackson – and how implicit biases might implicate us all in epistemic injustices today.

a. i. Testimonial Injustice

Epistemic injustice can occur when speakers attempt to provide knowledge, insights and understanding to other people. When a speaker attempts to impart knowledge, an assessment is made, by the audience, of whether the speaker is credible. Judgments of credibility involve judgments of the speaker's reliability and trustworthiness as a knower. The listener will (implicitly or explicitly) consider questions such as *Is she intellectually capable enough to have acquired or produced knowledge?* and *Is she trustworthy? Can I believe her?* Sometimes the resultant assessments are not based on the personal characteristics of the individual speaker, such as her track record, or her qualifications, or whether, for instance, she shows proficiency with the technical terms of a domain of knowledge. Instead they are based on prejudices about the individual speaker's social group. For example, a speaker might be deemed not to be intellectually capable because she is a woman and women are

thought to be intellectually inferior to men. Or a speaker might be deemed not to be trustworthy because she is black and black people are taken to be dishonest. This is part of what Collins has in mind when she says that 'power relations shape who is believed and why' (2000, 270). Sexist and racist stereotypes mean that some people are disempowered as knowers because of the stereotypes through which others see them.

More recently, the phenomenon whereby speakers are given less credibility than they deserve due to prejudice has been called *testimonial injustice* by Miranda Fricker (2007). Testimonial injustice occurs when speakers are treated as less reliable and trustworthy than they really are, and not believed when they should be, due to prejudice. We see both dimensions of this form of epistemic injustice in *Hidden Figures*. In the opening scene, Goble, Jackson and Vaughan are at the side of the road, broken down. Vaughan is under the engine, trying to fix it. Immediately on their guard, a police officer pulls up to enquire about the troublesome spot they are in. 'You being disrespectful?' the officer challenges Jackson, when she points out they didn't choose to break down there. His tone is overtly hostile – despite their deferential demeanor. The threatening tone of the exchange, however, rapidly changes when their ID reveals that they are employees of NASA. The default assumptions made by the police officer are revealed as he expresses surprise: 'I had no idea they hired...' '...there are quite a few women, working in the space program' - Vaughan saves them all the indignity of his racist utterance. Impressed by their employment, the police officer has nonetheless revealed his racist and sexist default assumptions; never did he suppose they could be mathematicians. Only the veneer of esteem that comes with doing the calculations to get the rockets into space for NASA suffices to undo the credibility deficit the police officer brings to their exchange.

There are lots of examples in the film of people not giving the protagonists the credibility they deserve. Fricker refers to these as credibility deficits. Testimonial injustice involving *credibility deficits* (Fricker 2007, 17) like this are harmful. Judging another person not to be credible, and not believing what they say, simply because of the social groups to which they belong is dehumanizing (Fricker 2015). Moreover, failing to accord people the credibility they deserve can damage the person denied credibility *and* other members of their community who do not appreciate and cannot make use of the knowledge that they provide. Others lose out on knowledge, too.

Thinking about epistemic injustice in terms of credibility also helps us to see the importance of markers of credibility – norms or indicators that help us to make accurate judgments of credibility. One such norm includes titles that indicate job roles and statuses – these help us to appropriately understand an individuals' level of expertise. Recall, the police officer swiftly changes his tone when the three women show their NASA passes – key markers of esteem and credibility. This helps us to see a further way in which one could suffer an epistemic injustice: by being denied the markers of credibility – as was Dorothy Vaughan as depicted in *Hidden Figures*. Vaughan has been undertaking work in effect supervising the team of black women working as human computers within NASA, but is denied promotion to supervisor, and thus denied a key marker of credibility. This is also makes it more likely that she will suffer testimonial justice in future, so paves the way for further harms.

a. ii. Testimonial injustice and implicit bias

Some people might optimistically think that many of these forms of epistemic injustice are rarer today, insofar as many of us reject the overt racism that characterized social relations in

the 60s. However, racist dynamics have persisted alongside the trend that has seen many people profess egalitarianism and a commitment to equal treatment. How can this be explained? Are people who say they care about equal treatment lying? Maybe. But recent research from psychology has provided a competing explanation: people who are sincere in professing a commitment to fair treatment may also have implicit biases. Implicit biases might be one of the mechanisms involved in perpetuating epistemic and other forms of injustice (other factors may include social structures, such as segregation, of the sort we mentioned above, and simply unjust and unequal access to material resources needed to gain and contribute knowledge).

Other chapters in this volume examine the nature and moral implications of our implicit biases, and how our automatic patterns of thought, feeling, and behavior can be informed by stereotypes. Some of these stereotypes and behaviors, as we will see below, concern people's capacities as knowers and testifiers. Thinking about implicit bias and epistemic injustice can also help us to refine our understanding of the injustices at issue.

Credibility involves being judged reliable as a knower, and trustworthy when one imparts knowledge. But we might have implicit biases that impact on these judgments; that is, we may have some automatic thoughts about the extent to which people are reliable and trustworthy. Some of the early studies on implicit bias suggested that people tend to be more ready to associate positive qualities with white people than with black people, especially when it comes to evaluating their competences. For example, in one study, (Dovidio & Gaertner 2000) individuals were asked to evaluate the qualifications and credentials of job applicants, and report back on how likely they were to recommend that the individual be hired. Sometimes the same materials would be identified as belonging to a black candidate and other times to a white candidate. When the applicant's materials indicated they were racialized black, and where there was room for discretionary judgement, the evaluations were less positive, and less strong hiring recommendations were made, than when the applicant was indicated as white. One explanation for these differential evaluations is that there are implicit biases at work here – automatic associations with positive qualities – that lead people to see white applicants as more suitable, competent candidates. Whilst the impact of implicit bias in any one case may be marginal, if these biases are pervasive – as they have been found to be – then they could be part of an explanation for widespread discrimination. This could be because implicit biases led the evaluators to regard the white applicants as more credible knowers. But maybe this has nothing to do with the applicants' knowledge relevant capacities. Maybe the evaluators have implicit preferences for white applicants, whether or not they (implicitly) think of white people as more reliable knowers? That could be – but other studies suggest that sometimes implicit biases concern specific evaluations of credibility and trustworthiness.

Some psychologists have looked specifically at whether implicit racial bias is related to the judgments of trustworthiness that people make. For example, Stanley et al (2011) first asked participants to take a race IAT. Then they asked participants to rate a series of faces for how trustworthy they appeared. They found that to the extent that individuals harbored stronger positive associations with white people than black people, they were also more likely to judge white faces as more trustworthy than black faces. This suggests that simply in virtue of being white, people are getting a credibility boost that is not afforded black people. Insofar as credibility judgements interact (cf. Anderson 2012, Coady 2017), this means that white people will be more likely to be believed than black people. Accordingly, the evidence from empirical psychology suggests that, even if people try to be fair-minded and treat

people equally, implicit bias might mean they are implicated in testimonial injustice. And these implicit biases may be widespread precisely because they reflect a social context in which there is unequal access to the markers of credibility, as we described earlier.

b. i. Epistemic appropriation

Testimonial injustice and credibility deficits can prevent the sharing of knowledge. However, epistemic injustice can also operate via credibility denial in a different way. In some cases, someone may produce knowledge, but the knowledge producer does not get credit for the idea. Their contribution is recognized, but that it is *their* contribution is not recognized. This has been called *epistemic appropriation* (Davis, 2018). For example, suppose a black student makes a contribution in class. Shortly after, a white student takes up the idea and the rest of the class discuss it as if it had not already been introduced. The important idea is recognized, but the student who first introduced it does not get due credit. Their idea has been appropriated.

As Davis explains, there can be voluntary epistemic appropriation, for example, someone from a group that is marginalized or stigmatized might allow her ideas to be published under another person's name because she does not believe that the ideas would be well-received if known to be hers. However, epistemic appropriation may often be involuntary; ideas can be adopted without recognition of who produced them. When epistemic appropriation occurs, the knowledge producer does not gain the boost to their credibility that they are due. Meanwhile those already in positions of power who are taken to be knowledge-producers can gain the benefit of undeserved recognition.

For example, in the film *Hidden Figures* Katherine Goble is responsible for authoring reports that are crucial to NASA's achievements in the Space Race. But she is denied the credit that she deserves for her contribution because her white male colleague, Paul Stafford, is named as sole author of the reports. Not only does Goble fail to receive the credit she is due; Stafford appropriates it instead. This harms her in various ways: if she had been named as author on the reports then she could have gained a boost to her status as a mathematician. She could consequently have been given opportunities to engage in other projects, further contributing to the knowledge in her field. In contrast, while her name is excluded from the reports her white male colleague has an unwarranted boost to his credibility—not only is he credited as author, he is wrongly credited as the sole author of the work, and taken to have been solely responsible for the knowledge contained therein.

Epistemic appropriation can occur over time also. For example, one of the projects of the book and film *Hidden Figures* is to ensure that credit is appropriately accorded to the black women who were central to the Space Race. Notwithstanding their crucial role, prior to the publication of the book and later release of the film their important epistemic contributions were not widely known. Yet recording their roles, as Margot Lee Shetterly does in the book, is a way of ensuring, over time, that epistemic appropriation does not persist, due credit is given, and that at least some of the wrongs of previous epistemic appropriation are rectified.

b. ii. Epistemic appropriation and implicit bias

Epistemic appropriation occurs when someone's ideas are recognized, but not properly attributed. Might implicit bias be involved in this? Patterns of implicit bias – gender and racial bias, say – might be implicated particularly if we think about the appropriation of

ideas at the level of social identities and widely held stereotypes or biases in a social context. An example of this concerns the kind of implicit stereotypes to do with trustworthiness that we talked about in relation to testimonial injustice. Patricia Hill Collins has also articulated the kinds of ‘controlling images’ through which black women in particular, in the US, may be seen: as nurturing and obedient ‘mammies’ (80); as unfeminine, emasculating ‘matriarchs’ (82); materialistic, domineering, but dependent ‘welfare queens’ (88); as professional but unfeminine ‘black ladies’ who take white jobs through affirmative action programs (89); or sexually aggressive ‘jezebels’ (90). These controlling images may be ones to which racist people knowingly subscribe; but many people reject these stereotypes. In the latter case, it may also be that the stereotypes are nonetheless held implicitly, influencing peoples’ judgments without them realizing or intending this to be the case. (As Goff & Kahn (2013) note, few studies on implicit bias have examined black women in particular, tending to focus on stereotypes that attach to white women or black men. This means that there is a gap in the psychological literature regarding the associations that black women might face). Why are black women seen through these kinds of stereotypes, when instead there are available to us inspirational stories such as those of the women who worked at NASA – such as Goble, Vaughan and Jackson? Collins traces the political utility of these stereotypes in entrenching the oppression and exploitation of black women.

In contrast, NASA, and aeronautical engineering, is widely stereotyped as the kind of endeavor characteristic of white men. Indeed, a study that examined the association between men versus women, and the sciences versus the arts, found that participants strongly associated men with the sciences (Nosek et al 2002).

As it was characterized earlier, appropriation concerned individual knowledge producers not being recognized. But we argue that we can think of the stereotyping of certain fields or subjects as a kind of collective epistemic appropriation. Activities in which black women made key contributions have come to be stereotyped as activities that are typically done by white men. The importance of that field is recognized, as is the ingenuity and scientific rigor of those working in it; the ideas as seen as fundamental to the advancement of human knowledge. But that the field was shaped *by black women*, that *their* ideas have been fundamental to the advancement of human knowledge, has been overlooked and ignored. Aeronautical engineers are stereotyped as white men. Instead of aeronautical engineers, black women are stereotyped with the controlling images that Collins articulates. This is a kind of collective epistemic appropriation. Insofar as implicit biases are implicated in these stereotypes and shape our patterns of what Charles Mills calls ‘collective remembering and amnesia’ (2007: 28-29) – remembering and forgetting *whose epistemic contributions they were*, in particular – we can see how people might, unintentionally, be implicated in this kind of epistemic injustice.

c. i. Epistemic exploitation

In a memorable set of scenes from *Hidden Figures*, Katherine Goble has to leave her office building and run half a mile across the NASA campus to reach the nearest available toilet, since the washrooms are segregated. She has asked the only other female in her office, a white administrator, where the toilet is and was told that the administrator did not know where *her* toilet was. In an intensive research environment in which people do not take breaks, it is quickly noticed that she is frequently away from her desk, and eventually her white male boss demands an explanation. In front of a large office full of white men (and one white woman), Goble is required to describe the difficulties and indignities that she has

encountered every time she has needed to go to the toilet. She articulates the frustration and humiliation she faces due to not having access to a toilet nearby and due to being made to drink from a coffee pot that none of her colleagues wish to even touch.

There are obvious obstacles to developing knowledge here – literally having less time to do so due to having to take longer bathroom breaks. But in addition, Goble is required to articulate the difficulties that she has faced within her work environment due to being a member of a stigmatized and marginalized group. She is placed in a situation in which she has to educate members of the dominant group about the inequities that she has faced. She suffers stress and embarrassment in doing so. Katherine's experience has many of the features commonly found in epistemic exploitation.

Epistemic exploitation occurs when members of a marginalized group are expected or required to educate members of privileged groups about injustices that are faced by those who share their social identity (Berenstain 2016; Davis 2016; Spivak 1999 and Audre Lorde 1990). Such educative work requires cognitive and emotional labor that is uncompensated (financially or otherwise), mentally draining, and time-consuming, taking their attention away from other activities that might have been rewarded. The efforts involved are not viewed as work, but people may face negative repercussions if they do not engage in them. For example, if a member of a marginalized group refuses to educate members of the privileged group, the people who made the request may be affronted. Moreover, stigmatized individuals may be viewed as confirming negative stereotypes about their groups, such as the stereotype that women are irrational, or that black people are uncooperative. These are all harms, and instances of injustice when social power explains why some people face repeated demands for burdensome cognitive work that goes unrecognized and unrewarded.

In the film *Hidden Figures*, Goble's explanation of her experiences leads to action, as her white male boss racially desegregates the toilets and removes the segregating labels from the coffee pots. Note, though, that this scene is not historically accurate, and also problematic: it perpetrates the myth of the 'white saviour' who came along to end racism. Whilst the film is of course entitled to fictional license, the problematic retelling of stories may be one of the reasons for which some authors have argued that there is a responsibility for the oppressed to educate others (Medina 2011) – such education is likely more reliable, and less likely to pander to problematic narratives whose aim is to make white people feel good.

Yet still, although members of marginalized groups are asked to provide information, the testimony provided may not be deemed credible precisely due to stereotypes about the marginalized groups to which they belong. This combination of epistemic exploitation and testimonial injustice explains why members of marginalized groups are required to repeatedly explain the injustices that they face. It can also result in individuals engaging in what Kristie Dotson calls *testimonial smothering* (Dotson 2011): truncating or silencing testimony to avoid the risks associated with not being properly listened to.

c. ii. Epistemic exploitation and implicit bias

How might implicit bias be involved in epistemic exploitation? In the face of systematic implicit bias, the onus on members of marginalized groups to educate others might be particularly heavy. First, where discrimination and inequalities result from implicit bias, they may be rationalized away, and so harder to recognize as inequalities or discrimination.

It is therefore more likely that a demand will be placed on members of marginalized groups to educate others about how specific judgments and actions resulting from implicit bias are unjust. Second, negative implicit attitudes towards members of marginalized groups might increase the chance that the cognitive and emotional labor that they contribute will be systematically undervalued, an important component of epistemic exploitation. Third, perpetrators of discrimination may be resistant to the idea that they are complicit in treating people unfairly, in part because it is difficult for them to notice that they have and are influenced by implicit bias. People may have misleading evidence – they think about their beliefs, and come to believe that they are fair-minded and do not treat people unfairly. They don't notice their implicit biases and the influence they have on their behaviors. In fact, there is some evidence that to the extent that people think they are being objective and are not influenced by bias, they are *more likely* to be biased (Uhlmann & Cohen 2007). In such cases, it may be harder for those trying to explain and educate about experiences of discrimination or injustice to have their testimony accepted. This might be so even if someone is explicitly *asked* to educate others; the hearer may find the testimony harder to believe if it does not cohere with their own evidence (namely, that they are objective and unbiased). So the fact that bias is *implicit*, and so hard to notice, may increase the likelihood of epistemic exploitation.

d. i. Hermeneutical injustice

The concepts that we have shape the way that we understand and communicate our experiences. There can be injustices surrounding whether people have access to and can utilize concepts and other conceptual resources (e.g. narratives, scripts) that capture and can be used to understand their experiences. Members of dominant groups can shape and unduly influence what concepts are widely available. For example, white, middle-aged, able-bodied, cisgender, and middle- or upper-class males have traditionally had access to positions and resources that enabled them to shape the concepts – the interpretive, or 'hermeneutical resources' – that are widely available. They have been the ones typically occupying positions of 'hermeneutical power'; they have been, for example, the politicians, journalists and educators – the people who have most influence over the concepts that are widely in use. Concepts that capture their experiences would be dominant. Meanwhile concepts that describe the experiences of people in positions of less power may not be widely understood – or even may not be available at all, until people undertake the cognitive and emotional work to try to articulate and describe aspects of their experience. For example, the concept of *sexual harassment* came to prominence in the 60s only after women worked collectively to understand and articulate the experiences that were making their participation in the workplace so difficult and so costly. (Note that this followed middle class white women's entry, post war, into the workforce. Black women and working class women - with fewer opportunities to shape the hermeneutical resources - had long been in the paid workforce, frequently with little choice about this, and had been subject to sexual harassment, assault and rape with no opportunity for recourse or redress.) Even once more prominent, the concept was not widely understood and not part of the shared conceptual resources (arguably the need for movements such as #metoo reflects the extent to which it is still not well understood). Indeed, it can be in the interests of the dominant group to exclude from the widely available conceptual resources concepts and conceptual resources that capture the experiences of marginalized groups (see Fricker 2007 for discussion of this example). This means that, for example, women who experienced sexual

harassment had – and still have – a hard time getting others to understand their experiences and what is wrong with it.

The notion of *hermeneutical injustice* has been introduced by Miranda Fricker (2007, chapter 7) to capture this phenomenon. She characterizes it as obtaining when individuals lack the concepts they need, due to a gap in the shared conceptual resources, which is in turn due to some groups having undue influence over the formation of those resources, with others having insufficient influence (*hermeneutical marginalization*).

As an instance of hermeneutical injustice, consider the experiences faced by Katherine Goble when she has to run half a mile to use the toilet. She is likely to be able to conceptualize this experience alongside numerous other difficulties she has faced: racist structures mean she faces obstacles meeting even her most basic needs, and are a feature of her experience she confronts regularly. Consider though the response of her white colleague after Goble has taken a necessarily lengthy bathroom break: “My God, where have you been? Have you finished yet?” One interpretation of the inability to comprehend is that her colleague lacks the conceptual resources to understand well the barriers that Goble faces. She just sees a long absence from the desk and an incomplete work package – she lacks the interpretive resources to conceptualize what Goble is experiencing in terms of racist social structures. Whilst Goble has the resources to make sense of her experiences, she is unable to communicate adequately about this with her colleague who lacks the relevant interpretive resources. Note, though, that Goble herself does not lack the resources to understand her experiences: she is hermeneutically marginalized, though, and therefore is unable to communicate with her colleagues about the racism she experiences, since they lack the relevant interpretive resources. This indicates we should expand the notion of hermeneutical injustice to include not only cases in which individuals lack resources to understand their own experience, but those cases in which they have the resources, but are unable to effectively communicate about them (e.g. to people in positions of power who could act to change those experiences) (Medina 2011; Pohlhaus 2012; Dotson 2012).

d. ii Hermeneutical injustice and implicit bias

Recall Goble’s difficulties in communicating with her white colleagues about the racist structures that hindered her work. Given what we have said about implicit biases being a contributor to testimonial injustice and epistemic exploitation in particular, one might think that the development of the concept of *implicit bias* is particularly helpful for dealing with this form of epistemic injustice. Having the concept of implicit bias helps people who think they are unbiased and committed to fair treatment realize that in fact they may not be. As such, it may help people to realize that despite their good intentions, and despite their values, they might nonetheless sometimes discriminate, and in particular perpetrate epistemic injustice. When they think of aeronautical engineers, they might think immediately of a white man, rather than a black woman. When hearing about experiences of discrimination, they might automatically think ‘that person is wrong, I don’t discriminate’. But if this person has the concept of *implicit bias* they might be better able to carefully reflect and notice that their automatic patterns of thought are biased. They might be more willing to listen when they are called out, because they acknowledge that their own perceptions of how they thought and acted may not be reliable (see Hahn et al (2014) for studies that suggest that when others prompt us to reflect, we can better notice our own biases). Moreover, having the concept of *implicit bias* could help to articulate widespread patterns of

discrimination and exclusion pervasively faced, even in contexts and interactions in which explicit prejudice does not prevail.

Having the concept of implicit bias can also prompt us to reflect on the kind of changes we need to make to society and institutions in order to make it less likely that implicit biases have a role (e.g. changes to hiring practices, reducing informal segregation, etc.). As such, having the concept of *implicit bias* in our shared conceptual resources can fill a gap in those resources. And once that gap is filled, it might help people recognize the ways that they and the social structures they inhabit are involved in discrimination, and motivate steps to address these injustices. Once again, however, the optimism expressed here needs to be qualified. The concept of implicit bias has been around for some time but there has been far from universal recognition of the phenomenon and the need for institutions to address the problems that follow from it. Implicit bias remains a contested concept in psychology (see Brownstein, this volume). This might be explained by the fact that recognition of implicit bias and its effects threatens the legitimacy of the power and status of privileged individuals. We should not expect the mere introduction of the concept of implicit bias to trigger actions to right epistemic and other wrongs. Instead, it is necessary to reflect carefully on how to use the concept effectively.

e. i. Contributory injustice

We introduced above the idea that some concepts may not be available to all, so that it is harder for some people to express or be understood in the claims that they make about their experiences. As we described, some concepts may exist in some groups, but may not (yet) be part of the dominant conceptual resources. *Contributory injustice*, as characterized by Kristie Dotson (2012), occurs when someone is willfully ignorant in using concepts that thwart others' abilities to *contribute* to the epistemic community. The injustice here is specifically in people being unable to contribute to the dominant shared interpretive resources (rather than in the existence of a gap that hinders understanding). Because some individuals and groups are unable to contribute to shared interpretive resources, it therefore becomes very likely that they will experience hermeneutical injustices of the sort we saw above. As we saw, there may be competing sets of concepts and conceptual resources (narratives, scripts, counter-mythologies) that exist among different social groups to explain the experiences of group members. The dominant conceptual resources may be structurally prejudiced, for example, by lacking concepts that some people need to make sense of, or communicate about, aspects of their oppression.

Consider the example of sexual harassment. The dominant conceptual scheme may conceptualize certain behavior as *banter* or *just a bit of fun*. So construed, it makes it harder for those who wish to use competing conceptual resources to capture the behavior – resources which capture the seriousness and the harm of the behavior they are experiencing – such as *sexual harassment*. The widespread use of competing resources (*banter*) poses obstacles for those using marginalized resources (*sexual harassment*) to express their experiences. Those who try to articulate their experiences in these terms, in particular to those who use the dominant conceptual resources, are thwarted in contributing knowledge and understanding. As we noted, when it is in the interests of people to ignore other important concepts, people are willfully ignorant – their ignorance is motivated by their interest in maintaining the status quo. Willful ignorance is helpfully characterized by Gaile Pohlhaus as “a willful refusal to acknowledge and to acquire the necessary tools for

knowing parts of the world" (Pohlhaus 2012, 729; see also Mills 1997 and Sullivan and Tuana 2007).

For example, in the film *Hidden Figures*, an interaction between Vivian Mitchell, the white supervisor of the black human computer team can be read in terms of contributory injustice. In one scene she, seemingly genuinely, declares that she has nothing against Dorothy Vaughan, who was at that time her subordinate – professionally, and socially, as a black woman in a racist social context. "I know. I know you probably believe that", Vaughan replies. Vaughan both believes that Mitchell doesn't have anything personally against her, *and* that despite her protestations, Mitchell really does harbor prejudice against her, due to her status as a black woman. One way to understand this is in terms of the competing concepts of prejudice that each of them have. Reading between the lines, it is likely that Mitchell has a conception of prejudice – one which figures in the dominant interpretive resources – that is characterized by race-based hatred or animosity. Since she doesn't hold such attitudes towards Vaughan, she isn't prejudiced against her (according to this conception of racial prejudice). Again reading between the lines, we might suggest that Vaughan holds a different conception of prejudice – one which does not depend on these kinds of mental states, but rather on the ideology that one accepts or assumes, as belied by one's behaviors. For example, that Mitchell accepts a workplace in which there are segregated job roles and assignments, bathrooms and coffee pots constitutes prejudice, and demonstrates Mitchell's subscription to a racist ideology, on this view, irrespective of whether she harbors animosity.

If the dominant conceptual resources suppose that racism must be characterized by hatred or animosity, then it will be harder for competing conceptions of racism – those which are a better fit for the commonplace experiences of racism that pervade societies structured by racist hierarchy – to be expressed and understood. The use of the more limited conception of racism could be a form of contributory injustice if those who use it do so in willful ignorance, namely, a willful refusal to gain the tools needed to understand properly the nature of racism in which whites were, and are, implicated. There is reason to suppose this is an accurate characterization of Mitchell, insofar as she willfully ignores important aspects of Vaughan's experiences or the concepts needed to make good sense of them, and insofar as she uses competing concepts that thwart Vaughan's (and others') abilities to get their experiences of racism well understood. In this instance, there is not only the epistemic harm (to members of marginalized groups) of being unable to contribute knowledge to shared understandings, but also the harm (to those who use the dominant conceptual resources) involved in people being poorly placed to understand and address all aspects of racism, which goes beyond the problematic mental states of racists.

e. ii. Contributory injustice and implicit bias

Contributory injustice occurs when people are marginalized and so unable to contribute to shared understandings, because the concepts they use to do so are not part of the shared resources. Others continue to use other shared resources that make it harder for marginalized people to make important contributions (recall the example of *sexual harassment/banter*). While we have suggested that having the concept of *implicit bias* is a helpful contribution to the shared resources, we here want also to raise a cautionary note. Using the concept of *implicit bias* might perpetrate contributory injustice if doing so makes it harder for marginalized people to contribute to the shared understandings. But why might

having the concept of *implicit bias* make this the case? Haven't we just argued that having the concept might make people *more likely* to recognize injustice?

The concern we want to raise is simply this. The concept of *implicit bias* comes from empirical research programs conducted by academic researchers. And, while it might be helpful in many ways, there is something problematic if people only heed the possibility that they are implicated in discrimination when academic researchers suggest as much. There are other sources of this knowledge: notably, the testimony of people who have experienced discrimination even when the discriminator claims not to be biased and professes values of fair treatment. Moreover, there is a pattern of evidence all pointing towards the same conclusion – that people discriminate without realizing it, or unwittingly stereotype, or express bias even though they try to be fair. These sources of evidence should also be given due weight. If we only heed these ideas when they come from academic researchers, this can make it harder for people in marginalized groups, who often do not have access to the platforms that academics have, and are denied access to some markers of credibility, to contribute. It entrenches the idea that certain ideas, expressed by certain people, in certain ways, are legitimate, while others are not.

So, in this context, only heeding the problems when academic researchers talk about *implicit bias* might make it harder for people from marginalized groups, who may use different concepts to capture their experiences of discrimination, to contribute to shared understandings. Whether this constitutes a form of contributory injustice might depend on whether the use of the concept *implicit bias* is done in willful ignorance, or whether willful ignorance is in fact an inessential component of the notion of contributory injustice. In any case, it is important to consider the research findings about implicit bias along other sources of evidence about the phenomena of and mechanisms involved in discrimination.

Implicit bias, epistemic injustice, and remedies

From the discussion above, we can see that addressing epistemic injustices requires a variety of different strategies. Sometimes it will require correcting stereotypes – including implicit stereotypes – that undermine credibility. If these are implicit stereotypes and biases, then creative strategies might be needed, since individual efforts alone may be insufficiently robust to secure change (Davidson and McHugh this volume, Madva this volume). Sometimes it will mean enabling people who are marginalized to access the markers of credibility. It might also involve revisiting what those markers of credibility are, and whether they are in fact good markers (should we rely on qualifications or prestige of schools, if not everyone has equal access to those markers of credibility?). It might involve ensuring that people get the recognition they are due, and ideas are properly credited to them; this can include challenging social stereotypes. It might involve recognizing unfair epistemic burdens, and taking on additional commitments to educate oneself. It might involve collective efforts to reflect on the shared concepts we have, and where they come from, and whether there are different or better concepts available to us. As we hope is clear from above, these kinds of strategies are best conceived as collective projects – shaping our shared understanding of what indicates credibility; challenging the social stereotypes that foster appropriation of ideas; shaping our shared interpretive or hermeneutical resources.

Whilst there are some things that we can do as individuals – we can educate ourselves, we can try to be better listeners, we can try to reflect on our automatic judgments – these will be of limited efficacy against a backdrop of systemic biases and history of structures of exclusion. Many of the changes needed are social changes – changes to

structures that prevent people accessing knowledge or communicating it effectively (see Ayala-Lopez this volume; Madva this volume). Whilst formal segregation has ended, de facto segregation – of housing, jobs, and social groups – still persists, meaning that there persist barriers to accessing knowledge, accessing markers of credibility, accessing and shaping the relevant conceptual resources for making sense of injustices. Changing social structures is not just a matter of justice, but also a matter of removing obstacles to knowledge, and opportunities to contribute to knowledge (Anderson 2010). Remedying epistemic injustice – including epistemic injustices due to implicit biases – will require changes to what and who are on the curricula that we learn, and how that affects whether we implicitly associate, for example, aeronautical engineering with white men or black women; changes to our shared concepts, which requires what Kristie Dotson calls ‘transconceptual communication’ – the ability to interact with people across social boundaries and try to understand the different concepts people use. This suggests that in addressing these various dimensions of epistemic injustice, we need to think about what we can do as individuals, but also as individuals who, with others, can bring about broader social change.

Where does this leave us with regards to our answer to the opening question of how knowledge and social power relate to each other? We have seen that there are many important ways that the two interrelate. Unjust social systems can prevent knowledge from being produced, acknowledged and acquired through phenomena like testimonial injustice, hermeneutical injustice and epistemic appropriation. And where some individuals in society lack knowledge about the social experiences of members of marginalized groups this can produce injustices like contributory injustice and epistemic exploitation. Knowledge and power are deeply intertwined.

Discussion questions:

1. Could it ever be just to give someone *more* credibility than they deserve? Why or why not?
2. What sorts of markers do we use to indicate credibility? Are these likely to be good, or reliable markers? Are there ways in which these markers might reflect, or entrench, injustices?
3. Consider cases where someone voluntarily allows her ideas to be appropriated, knowing that her ideas will be better received if reported by someone else. Does the fact that it is voluntary make it unproblematic? Why or why not?
4. Do individuals from marginalized groups have responsibilities to educate others about aspects of oppression? Or is doing so always epistemically exploitative?
5. Consider notions that have been recently developed: *mansplaining*, *manspreading*, *he-peating*. To what extent is it legitimate to think of these notions as filling a gap in the hermeneutical resources? Would the absence of such concepts have been a hermeneutical injustice?

6. Is *implicit bias* a useful concept for identifying instances of epistemic injustice? Or might its use sometimes – or always – perpetrate contributory injustice?

References

- Anderson, E. (2012). Epistemic justice as a virtue of social institutions. *Social epistemology*, 26(2), 163-173.
- Anderson, Elizabeth. *The imperative of integration*. Princeton University Press, 2010.
- Berenstain, N. (2016). Epistemic Exploitation, *Ergo*.
- Coady, D. (2017). Epistemic injustice as distributive injustice. *The Routledge Handbook of Epistemic Injustice*. Routledge, pp. 79-86.
- Collins, P. H. (2000). *Black feminist thought. Knowledge, Consciousness, and the Politics of Empowerment*, second edition. New York: Routledge.
- Davis, E. (2016). Typecasts, tokens, and spokespersons: a Case for Credibility Excess as testimonial injustice. *Hypatia*, 31(3), 485-501.
- Davis, E. (2018). On Epistemic Appropriation. *Ethics*, 128(4), 702-727.
- Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia*, 26(2), 236-257.
- Dotson, K. (2012). A cautionary tale: On limiting epistemic oppression. *Frontiers: A Journal of Women Studies*, 33(1), 24-47.
- Dovidio, J. F., & Gaertner, S. L. (2000). "Aversive racism and selection decisions: 1989 and 1999". *Psychological Science*, 11: 319-323.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Fricker, M. (2015). Epistemic contribution as a central human capability. In George Hull (ed.) *The Equal Society: Essays on equality in theory and practice*. Lanham, MD.: Lexington.
- Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, 143(3), 1369-1392
- Goff, P. A., and Kahn, K.B. How psychological science impedes intersectional thinking. *Du Bois Review: Social Science Research on Race* 10.2 (2013): 365-384.
- Hidden Figures*, 20th Century Fox, <https://www.foxmovies.com/movies/hidden-figures>
- Mason, R. (2011). Two kinds of unknowing. *Hypatia*, 26(2), 294-307.

Medina, J. (2011). The relevance of credibility excess in a proportional view of epistemic injustice: Differential epistemic authority and the social imaginary. *Social Epistemology*, 25(1), 15-35.

Melfi, T. (2016). *Hidden Figures*. United States: Fox 2000 Pictures, Chernin Entertainment and Levantine Films.

Mills, C. (1997). *The Racial Contract*. Ithaca, N.Y.: Cornell University Press.

Mills, C. (1998). *Blackness visible: Essays on philosophy and race*. Ithaca, N.Y.: Cornell

Mills, C. (2007). White ignorance. In S Sullivan and N Tuana eds. *Race and epistemologies of ignorance*, New York: State University of New York Press, 26-31.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1), 101.

Pohlhaus Jr, G. (2012). Relational knowing and epistemic injustice: Toward a theory of willful hermeneutical ignorance. *Hypatia*, 27(4), 715-735.

Shetterly, M.L. (2016). *Hidden Figures: The Untold Story of the African American Women Who Helped Win the Space Race*. New York: William Morrow and Company.

Spivak, G. (1999). *A critique of postcolonial reason: Toward a history of the vanishing present* (Cambridge, Mass.: Harvard University Press).

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences*, 108(19), 7710-7715.

Sullivan, S. & Tuana, N (2007). *Race and epistemologies of ignorance*. New York: State University of New York Press.

Tuana, N., & Sullivan, S. (2006). Introduction: Feminist epistemologies of ignorance. *Hypatia*, 21(3), vii-ix

Uhlmann, E. L., & Cohen, G. L. (2007). "I think it, therefore it's true": Effects of self-perceived objectivity on hiring discrimination. *Organizational Behavior and Human Decision Processes*, 104(2), 207-223.

Key words: racial bias; gender bias; epistemic exploitation; epistemic appropriation; testimonial injustice; hermeneutical injustice; contributory injustice; structural change.