



Social Epistemology

A Journal of Knowledge, Culture and Policy

ISSN: 0269-1728 (Print) 1464-5297 (Online) Journal homepage: <http://www.tandfonline.com/loi/tsep20>

Sham Epistemic Authority and Implicit Racial Bias

Charles Lassiter

To cite this article: Charles Lassiter (2019): Sham Epistemic Authority and Implicit Racial Bias, Social Epistemology, DOI: [10.1080/02691728.2018.1552995](https://doi.org/10.1080/02691728.2018.1552995)

To link to this article: <https://doi.org/10.1080/02691728.2018.1552995>



Published online: 11 Jan 2019.



Submit your article to this journal [↗](#)



Article views: 7



View Crossmark data [↗](#)



Sham Epistemic Authority and Implicit Racial Bias

Charles Lassiter

Department of Philosophy, Gonzaga University, Spokane, WA, USA

ABSTRACT

In this paper, I argue for a way to evaluate beliefs grounded in the testimony of sham epistemic authorities and then apply that principle to beliefs tokened as a consequence of implicit biases. As a result, such implicitly biased beliefs are correctly evaluated as praiseworthy by their possessors, even though the belief is epistemically defective.

KEYWORDS

Implicit bias; epistemic authority; testimony

1. Introduction

After a May 2017 article on implicit bias ran in *The Atlantic*, readers wrote in with stories describing times they witnessed or were subjected to discrimination caused by implicit biases. Here is one representative case:

Another reader recalled a bank manager's two reasons for why he brought a white man with a high-school degree into a management-training program, but not the head teller – a black woman who'd graduated college: "He reminded me of myself when I was just starting out." And, even more damning (but still totally unconscious): "He just looked like a banker."¹

This one is worth focusing on for a few reasons. First, the bank manager made the wrong choice: the candidate with experience and a college degree is likely to be superior to one with less experience and no degree. Second, while implicit racism might be widespread, explicit racism is less so. Presuming the manager is doing his level best is not being overly charitable. A perfectly reasonable gloss on the account is that the manager's belief is formed, in part, by implicit racial biases: a cognitive construct not accessible upon reflection that influences belief-formation processes. His belief is an implicitly biased belief; his judgment an implicitly biased judgment.²

The discovery that forces outside the agent's awareness motivate belief segue naturally into concerns about responsibility (cf. Holroyd, Scaife, and Stafford 2017). Discussion of epistemic responsibility over implicitly biased judgments run the gamut from absolution of blameworthiness (provided the agent is unaware of the implicit bias, e.g. Saul 2013; Levy 2017), to conditional blameworthiness (e.g. Washington and Kelly 2016) to nearly unconditional blameworthiness (e.g. Holroyd 2012, 2015). Assessments of blameworthiness hinge on awareness, or potential awareness of the implicit bias. For example, Saul (2013) argues that agents harboring implicit biases are not blameworthy for their biased judgments because (i) they are not aware of the unconscious bias and its influence, and (ii) biased agents cannot help have these sorts of implicit biases from living in a discriminatory culture. Saul continues that once an agent *is* aware that they harbor implicit biases against particular groups, they become culpable for doing something about it. In contrast to Holroyd (2015) argues that even if biased agents are not aware, they *ought* to be. Awareness of bias is a morally relevant feature of an action, Holroyd argues, and is therefore something of which

agents should be aware. Furthermore, not only *ought* agents be aware of their biases, experimental data suggests that agents *can* have this kind of awareness.

These are just two positions among several that have explored in the literature.³ And there are important and subtle differences among philosophers thinking about responsibility (both epistemic and moral) for implicitly biased beliefs. Even so, all share an important feature. None entertain the justificatory role of an implicit bias's etiology in assessments of blameworthiness. That is, social roles and structures factor in causally but not normatively. Let us consider an account in which social roles and structures feature prominently. Washington and Kelly (2016); Holroyd (2015); and Holroyd and Kelly (2016), develop externalist accounts of responsibility assessments in which social roles and structures are foundational.⁴ They argue that implicitly biased agents are morally responsible for actions motivated by implicit biases because agents have ecological control over their actions: agents are capable of manipulating their environments in a way that reduces the influence of implicit biases.⁵ In manipulating environments, agents shape their cognitive processes and character. By the authors' lights, the account is externalist because 'not all of the knowledge relevant to moral responsibility and exculpation need be "in the head" of the agent.'⁶ Because cognitive processes include the environment in some way, and manipulating the environment is within the control of the agent, agents can indirectly shape their own cognitive processes. The conclusion is that agents not only *ought* to subdue their implicit biases, but they *can* also through ecological control of their own cognitive processes.

But notice the role of the environment on this account: it plays a *causal* role in the development and maintenance of implicit biases. For Holroyd and colleagues, the social environment is important because it causally impacts an agent's character, and it is character which grounds normative assessments of an agent's beliefs and actions: an agent's actions are morally wrong because they flow from (unconsciously) discriminatory character. This is in contrast to the social turn in epistemology which sees social relationships, statuses, and structures as grounds for normative assessments of belief. In the epistemology of testimony, for example, it is not enough that one come by some bit of true information. *How* one comes by the information is just as important in determining the epistemic status of the resulting belief. Believing that *p* because I overheard someone in another room is different from believing that *p* because a trustworthy and reliable source confidently asserted to me that *p*. In both cases, I might believe *p* but I'm only justified in the latter case. But the normative relevance of a belief's etiology – what separates the good testimonial belief from the bad – has been missing from recent discussions of implicitly biased belief. This, I think, is a mistake for it fails to appreciate the role social forces can play in the justification – and not just the formation – of a biased belief.

I will argue that individuals are correct in positively appraising their implicitly biased beliefs upon reflection. The main premise for the argument is a non-controversial claim – **Belief from Sham Authority** – which says that beliefs acquired from such merely apparent authorities seem to the believer upon reflection to be reasonable but are nonetheless epistemically defective. **Belief from Sham Authority** reflects that the inquirer-authority social relationship has normative, and not just causal, import in evaluations of implicitly biased beliefs. How one should evaluate one's beliefs depends, in part, on who one takes to be an authority.

It is worth noting at this point what I am *not* arguing; first, that the biased agent's belief is true. Go back to the case of the bank manager. Reflecting on the evidence, he ought to have chosen otherwise. Even so, we're after how the manager ought to assess his own belief, not whether his belief is true. To say that the manager ought not assess his belief positively because reflection on the evidence shows otherwise is the epistemic equivalent of Monday morning quarterbacking: in hindsight, the right decision is always clear. Second, I'm not arguing about whether there are negative moral and epistemic consequences for others based on the biased agents' judgments. I think it is clear that there are, and that negative consequences can result from judgments one believes to be well-grounded. Finally, I am not arguing whether there is anything biased agents can

do to correct for their biases. The data show there are and we ought to undertake those actions (cf. Lai, Hoffman, and Nosek 2013).

How implicitly biased agents appraise their own belief-formation processes is an important issue. Implicit biases influence in subterranean ways, and their effects are unnoticed. In the best of circumstances, among reflective people who possess a concept of implicit bias, discriminatory contents can influence belief formation. But contrary to the claims of Washington and Kelly (2016), the concept of implicit bias is likely foreign for the majority of Americans (see appendix 1 for data and argument). For folks who do not possess this concept or do not think that they themselves are subject to such biases, philosophers ought to be concerned with how they think about the evidence they possess in support of their beliefs. It is difficult to blame people who have never heard of implicit biases or think research into the unconscious is so much hokum. It's not *prima facie* obvious that people unexposed to current research in social psychology bear heavy responsibility for their implicit biases. And if agents doing their level-best find their beliefs well-formed, then there is little motivation for them to change their minds. Folks entertaining implicitly biased beliefs are like Aristotle's sea captain: he chucks cargo overboard in a storm to prevent the boat from sinking. While jettisoning cargo is a bad thing, we evaluators of his action understand why he did it and even that the captain sees it as a good thing. In the same way, people's implicitly biased beliefs are bad because they rely unconsciously on stereotypes and inaccurate information about a population. But just as the sea captain's actions are by his own lights praiseworthy, so too are the biased person's judgments praiseworthy by his own lights, or so I will argue.

Of particular interest at present is *how* agents' implicit biases are maintained and reinforced over time. There are many ways, but I will focus on one in particular: biases that are grounded in testimony from authority. Authorities whose testimony grounds implicitly biased beliefs are *sham epistemic authorities* (or 'shams' for short). I will have quite a bit more to say about sham epistemic authorities, but here's a teaser. Sham epistemic authorities are individuals who look and sound as if they are epistemic superiors with respect to some domain, but the appearance is illusory. They are all overtone and no substance. A case of sham epistemic authority is the expert who speaks authoritatively in areas outside their domain of expertise (cf. Ballantyne [forthcoming](#)). Linus Pauling, for example, won two Nobel Prizes but advocated megadoses of vitamin C as a way to cure all manner of ailments, including cancer. Pauling – clearly an authority in his own right – stepped beyond the bounds of his expertise. When advocating vitamins, Pauling was a sham epistemic authority.

Here is how the paper will unfold. First, I will argue for the previously mentioned principle about evaluating belief grounded in testimony from sham authorities. After this, I will spend some time discussing the nature of sham authorities. In §4, I will offer a set of psychological mechanisms by which information is transmitted from speakers to hearers which amplifies the epistemic effects of implicit biases (more on these effects below).⁷ In §5, with the normative and empirical accounts developed, I will apply the principle defended in §2 to cases of implicitly biased beliefs. In §6, I respond to some objections.

2. Beliefs from Sham Authorities (BSA)

We begin with the uncontroversial claim about evaluating belief grounded in testimony from shams:

Beliefs from Sham Authorities (BSA) beliefs grounded in testimony from sham authorities are reasonable to the inquirer on reflection but are epistemically defective.

BSA claims that the believer, on rational reflection, finds the belief reasonable to hold when it is grounded in testimony from the sham authority, with testimony in this case including what hearers are 'invited' to believe (cf. Adler 1997). It does not specify the nature of the epistemic defect, for there is no need to do so.⁸ In fact, **BSA** requires only the possibility that a belief can be rational

upon reflection as well as defective in some way.⁹ **BSA** develops this possibility by stating a particular etiology for the belief: testimony from a sham epistemic authority.

Why endorse **BSA**? There are a few suggestive lines of evidence. First, **BSA** is one development of the distinction between a belief's rationality and its warrant (DePaul 1993).¹⁰ A belief is rational when a believer, on reflection, would take the belief to be true. A belief's warrant, on the other hand, is a matter of its fittingness to experience. So a belief that there is coffee in the mug has warrant provided all my sensory experiences suggest just that (and there is no evidence that my senses are untrustworthy in this instance). Warrant for DePaul, in other words, is an external and objective matter. **BSA** states that a belief from a sham authority can be rational and yet lack warrant.

Additionally, **BSA** isn't new; Aristotle had similar insights. In classifying his four moral character types, Aristotle describes the vicious agent as someone who believes that he is doing well even though he is not. The vicious character's understanding of the good is misdirected; active pursuit of wrongdoing is (mis)understood as enacting virtue. Here, Aristotle's insight is that moral agents might deceive themselves about pursuit of the good. They really might be doing their level best, and really might come up short. And it might be obvious that your best isn't good enough in calculus or logic, but it's less clear with conflicting moral goods or discerning who has your best interests at heart. When you do come up short, despite doing your level best, your beliefs are still defective. **BSA** captures these insights.¹¹

Applying **BSA** to concrete cases furthermore gives intuitively right answers when evaluating the epistemic quality of beliefs from shams (much more on these folks in §4). It expresses both that the inquirer's belief is faulty but also that the inquirer is correct in evaluating her belief as rational. Evidence comes from comparing cases of testimony by genuine and sham authorities:

Real Deal: John has an unexpected but pleasant conversation with a professor of classical history at a coffee shop. In the course of their conversation, the professor gives off a number of subtle cues about being a professor. She kvetches about classes, gripes about grading, and complains about committee work. She also asserts many claims about ancient Greece and Rome, some of which John knows to be true and others that he has no evidence one way or the other. The professor asserts, "Pirates on the Aegean Sea captured Julius Caesar when he was in his 20s. They initially asked for 20 pieces of silver as ransom, but Caesar insisted they ask for 50." John comes to hold this belief.

Con Artist: Susan has an unexpected but pleasant conversation with a person at a coffee shop. This person kvetches about classes, gripes about grading, and complains about committee work. The person asserts many claims about Greece and Rome, some of which Susan knows to be true and others that she has no evidence one way or the other. The person asserts, 'Pirates on the Mediterranean Sea captured Julius Caesar when he was in his 30s. They initially asked for 30 pieces of silver as ransom, but Caesar insisted they ask for 90.' The person, unknown to Susan, is a con artist. This con artist, to keep up the ruse repeats things that he vaguely recalls overhearing on *The History Channel*. Among the things asserted is the claim about Caesar's ransom. Susan comes to hold this belief about Caesar.

Reflections on these cases offer reasons to endorse **BSA**. First off, both John and Susan are acting reasonably in holding their respective beliefs about Caesar. How? Two components are important. First, the belief isn't obviously absurd; it doesn't fly in the face of other propositions for which (we can suppose) they have good evidence. Second, both the professor and the con artist give off signals that establish their credibility as epistemic authorities. They assert at least some truths in the course of the conversation. They speak confidently about history. They complain about the kinds of things that (we can suppose) John and Susan know professors to complain about. Reflection on this evidence for their belief about Caesar's ransom suggests that it's

reasonable to hold the belief. What are the conditions for the beliefs being ‘reasonable’ to hold? A rough-and-ready notion will do: reflective thinkers in the same context as Susan and John would come to hold the same belief (cf. Zagzebski 2017).¹² How reflective thinkers might do this depends on your preferred theory of warrant. The following are all consistent with the rough-and-ready notion of reasonability: exercises of properly functioning mechanisms or performances of intellectual virtues or careful evaluation of the evidence.

But a concern lurks nearby: why should reflection on the social cues of the speaker’s authority confirm that the speaker is an authority, making the belief reasonable to hold? The answer is because social signals are all we have to go on, most days. I’m going to presume that my students have never checked that I actually have a PhD in philosophy from an accredited institution (I never did when I was an undergrad). Nor do I check my physician’s standing with the American Medical Association before seeing her about a sore throat. By and large, we rely on social cues to indicate social roles, including expertise. Consider a study showing that 95% of hospital nurses complied with a doctor’s orders to deliver an amount of medicine that would kill the patient. (The nurses were stopped on their way to deliver the medicine and debriefed about the experiment.)¹³ Nurses in this case relied on a social regularity: doctors having expert medical and pharmaceutical knowledge. There’s simply not time or world enough to check everyone’s credentials. Of practical necessity, we depend on the correspondence between social cues and authority. Typically, this correspondence is reliable because we have social institutions to preserve the correspondence. Such is the case with Susan and John. They rely on widely recognized markers of authority to signal the speaker’s epistemic standing, which is a perfectly reasonable thing to do. By analogy, if a SWAT team swarms Port Authority Bus Terminal and one shouts at me to stay out of the subway, it is unreasonable (and probably imprudent) to epistemically require that I ask for their police badge numbers to ensure I am not being duped.

But even though both Susan and John are rational in believing the false fact about Caesar, Susan’s belief (but not John’s) is defective. Why? Our con artist isn’t an authority but is only impersonating one. Impersonating an authority doesn’t make one authoritative. Accepting the con artist as an authority would be like getting medical advice from an actor who plays a doctor on TV. Both can sound convincing, but you’re putting yourself in epistemic peril (if not mortal!) by believing the speaker. So even though Susan’s and John’s beliefs are reasonable upon reflection to hold, only Susan’s belief is defective.¹⁴

I’ve argued that **BSA** is a reasonable way to evaluate beliefs formed on the basis of sham authorities. We turn now to examining the nature of sham authority in greater detail.

3. Sham Epistemic Authorities

BSA evaluates beliefs grounded in testimony from sham epistemic authorities. But what exactly is a sham authority? We begin with an example: Mac from John Steinbeck’s novel *In Dubious Battle*. Mac instigates a strike of migrant fruit pickers against the landowners over a recent wage cut. Though Mac, along with his comrade Jim, is leading the strike, Mac knows he’s an outsider to that migrant community and needs a way in. The opportunity presents itself when London, one of the workers, reports that his daughter-in-law Lisa is ready to give birth. But there’s an older woman in the camp who, we are lead to believe, knows quite a lot about birthing babies. That’s where we pick up the story:

Mac turned to the old woman. ‘You a midwife?’

She scratched the back of her wrinkled hands and looked vacantly up at him, but she didn’t answer. ‘I asked if you was a midwife?’ he cried.

‘No – but I’ve took one or two babies in my life’

Mac reached down and picked up one of her hands and held the lighted candle close to it. The nails were long and broken and dirty, and the hands were bluish-grey. 'You've took some dead ones, then,' he said. 'What was you goin' to use for cloths?'

The old woman pointed to a pile of newspapers. 'Lisa ain't had but two pains,' she whined. 'We got papers to catch the mess.'

London leaned forward, his mouth slightly open with attention, his eyes searching Mac's eyes. The tonsure shone in the candle-light. He corroborated the old woman. 'Lisa had two pains, just finished one.'

Mac made a little gesture towards the outside with his head. He went out through the tent-flap and London and Jim followed him. 'Listen,' he said to London, 'you seen them hands. The baby might live if he's grabbed with hands like that, but the girl don't stand a hell of a chance. You better kick that old girl out.' (Steinbeck 1936, 55)

Fortunately, Lisa delivers her baby successfully with Mac's help. Jim later marvels at Mac's work, adding that he never knew Mac worked at a hospital. Mac then admits he has absolutely no medical training. 'The only thing I knew,' Mac tells us, 'was that it was a good idea to be clean.' He adds, 'That old woman knew lots more than I did. I think she knew it, too.'

There is a lot to consider in this example, but let's focus for now on Mac's status as a sham authority. He knows nothing about birthing babies, crafting his utterances and actions to give the appearance of authority. For example, he points out that the old woman is not a midwife and has lost a few babies. Her fingernails are dirty. He cross-examines her on the materials she has available ('what are you goin' to use for cloths?'). He study's Lisa's face and appears to discern evidence not available to the average eye, which confirms what is mutually known (that Lisa has had two contractions). On this last point, he does add in information that no one has yet mentioned – that one of her contractions had 'just finished.' But none of this is insightful. Of course an old woman who has no medical training has lost a few babies in childbirth in rural 1930s America. And of course the old woman has dirty fingernails; she's living in a camp of migrant fruit pickers in California. The most Mac seems to offer as evidence of authority is his assertion that Lisa has 'just finished' having a contraction. But even that's unconvincing: Lisa is in the throes of giving birth and 'just finished' is procrustean enough to fit whatever length of time is needed. But still, all this is enough to convince London that Mac knows what he is doing, more than the old woman. Mac is a sham. He has the appearance of greater authority about childbirth but most certainly does not.¹⁵

Readers might notice a similarity between the concepts of sham authority and Fricker's (2007) notion of *credibility excess*. Credibility excess is attribution to a speaker of significantly more credibility than he or she is due. Credibility excesses are often due to unjust socio-historical circumstances (cf. Medina 2010). Consider Mayella Ewell. She is the White woman who accuses Tom Robinson of rape in Harper Lee's *To Kill a Mockingbird* and is given more credibility than deserved because she's white (Medina 2010). Because she is White, her word is thought of as carrying more weight than a Black man's. She enjoys a credibility excess, while Tom suffers a credibility deficit.

While these are similar notions, they are nonetheless different. An excess of credibility does not make a sham. Mayella Ewell, for example, enjoys a credibility excess in her testimony against Tom Robinson. In virtue of this, her testimony is thought of by her hearers as more believable than Tom's. However Maya Ewell is not an authority. There is no expertise she possesses on the matter in which she is believed. In cases of sham authority, a speaker co-opts signals of authority. These signals typically pick out authorities, but they can be deployed by folks who are falsely believed to be in the know. Authorities that are picked out by these signals – also called 'indicator properties' – are *genuine epistemic authorities* (cf. Fricker 1998). People feigning authority by means of indicator properties are *sham epistemic authorities*.

What are indicator properties? They are socially salient properties that are typically reliable for identifying authorities. Their reliability comes from their track record: people who look like *this* or act like *that* typically have been good guides (or regarded as good guides) for learning thus and so. If I'm in a room of people and should happen to have a question about Catholic doctrine, odds are asking the man in the Roman collar is my best bet.

At this point, readers might wonder whether Goldman's (2001) sources of evidence used for determining trustworthy experts might have greater epistemic relevance than indicator properties in forming beliefs about authorities.¹⁶ While Goldman's evidence for expertise and indicator properties functions in similar ways, there are important differences between authorities and experts that render Goldman's evidence unsuitable for distinguishing between genuine and sham authority. First, there are authorities who are not experts in Goldman's sense. Accordingly, they are not picked out as experts in Goldman's sense. I might seek my parents' advice, for example, when I have a colicky baby who just won't sleep. Although my parents haven't been vetted by the relevant communities of experts, it is still reasonable for me to take their advice. So for authorities who are well-placed to give an informed opinion but not in areas where peer review is the norm for establishing this, indicator properties step in to help guide inquirers. Second, Goldman's properties are veritistic: they are truth-oriented. For example, if expert E is positively appraised by meta-experts, presumably those meta-experts are gauging E's reasoning abilities, her methods of data collection etc. The sorts of things for which E is judged are things where excellent performance brings one closer to truth. Indicator properties are not veritistic properties. They are socially-salient properties. But why not make indicator properties veritistic properties? There are two reasons. First, since not all authorities are experts in Goldman's sense, veritistic properties are useless for identifying such epistemic authorities (cf. Zagzebski 2012). Second, indicator properties better reflect the hurly-burly of human life. In some cases, we are able to reflect on evidence supporting an expert's standing. In many cases, we are not. Here, there are two kinds of cases. I might not be able to reflect on an authority's standing because of time pressures. For example, I might have to trust the man dressed as a conductor about which train I'm to get on if I have mere seconds to choose. In other cases, I am inducted into certain practices in which I look to certain kinds of figures for advice (cf. Rogoff 1990). As a result of my upbringing, for example, I might be more inclined to seek a chiropractor's advice for a persistent back-ache than a surgeon's. Or I might be disinclined to ask a police officer for directions if other options are available to me. In these kinds of cases, I do not always reflect on the standing of the authority because my upbringing has inclined me towards clinging to some authorities over others. Why do I incline towards some authorities over others? We, like Wittgenstein, might be inclined to say, 'This is simply what I do.'

There's an important point in the neighborhood: indicator properties, once identified, can be co-opted. Consider the rise of political pundits. Their job is to help interpret and understand the news but also to provide entertainment. The art of punditry, however, more often than not elevates looking smart over being smart. This is precisely what is taught at pundit schools, where ordinary folks learn how to be pundits. In case you're interested in a career change, top pundit tips include: using actionable words and phrases (e.g. referring to a political opponent as a 'dangerous psychotic'), insulting interlocutors, interrupting other speakers, and having a range of go-to phrases to shift conversation in the direction you want. ('Look, I think the real issue here is...') The trick? Appear authoritative in a political debate but without fussing over details like history or policy.¹⁷ Tricks like interrupting people, talking over others, and talking for longer than others all correlate with perceived dominance (Mast 2002). Exploiting these devices is the stock-in-trade of pundit schools, which then crank out strings of sham authorities.

How might indicator properties fail to pick out genuine authorities? Here are two ways this can happen. The first is a historical account from Fricker (1998). Social power has traditionally been tightly linked with access to educational institutions. Higher social power has offered access to more elite institutions with greater resources. By contrast, lower social power has offered access only to institutions with overstretched and insufficient resources. To find people with access to elite

institutions, you could just look for visually available markers of social power. This connection enables markers for social power to run proxy as markers for epistemic virtues like competence and trustworthiness. And over time, the markers and what they track can come apart. The 17th-century English gentleman is a case study (Shapin 1994; discussed in Fricker 1998). He was afforded competence in all matters and was regarded as trustworthy. Why? He had access to elite institutions. His testimony could be trusted since no one could financially incentivize bearing false witness. By contrast, non-gentle persons and women were missing these social markers, so they were believed to lack many epistemic virtues. Notably, they were not well-educated, and they were poor so their testimony was thought to be buyable for the right price. Initially, wealthy, white men were genuine authorities because they were educated. The social markers of being white, wealthy, and male came to function as indicator properties because of the initially high correlation between the markers and what they pick out. But in functioning as indicator properties, agents might possess the indicators without actually being an authority. So while whiteness, wealth, and maleness were initially only correlated with being an authority, they came to function as indicators, thereby enabling some to become shams.¹⁸

The second comes from developments in network science. Network science studies the emergence of networks and the flow of information across them. Networks consist of nodes and their relations. The simplest of networks is two connected nodes. Adding nodes and connecting them in different ways makes information flow more or less efficient. For instance, if we add a third and fourth node and connect them linearly, then we have a small network in which information can get to everyone but not in the most efficient way possible (e.g. Albert tells Beatrice, who tells Carlo, who tells Doreen). If everyone knows everyone in the network, then information is known by all much more quickly (e.g. Albert tells Beatrice, Carlo and Doreen).¹⁹

The target network structure of interest right now is a scale-free network.²⁰ In these networks, a few hubs are connected to many individuals, and many individuals are attached to just a few. There are many examples in the social world: co-citation networks in philosophy, the Twitter social media network, the World Wide Web, and directly connected airports. Now imagine an epistemic community structured as a scale-free network. A few folks are connected to many others, while many are connected to a few. One interpretation of the community structure is this. Connections indicate who goes to whom for advice. In a scale-free community, many people go to a handful of others for advice. Those in the handful are epistemic authorities. They are trusted by the members of their community for delivering advice. So in our interpretation of an epistemic community organized as a scale-free network, high degrees of connectivity indicate authority.

However someone might be highly connected for reasons independent of their epistemic excellence. Sometimes the connections are for other reasons: wit, physical attractiveness or wealth, for example. Sometimes there might not be reasons at all: scale-free networks can form without choices by rational agents. In fact, Barabási and Albert (1999) show how this can happen. They call the mechanism 'preferential attachment.' Start with a handful of nodes that are connected to one another. Introduce a new node. That new node picks one of the links at random and then one end of one link at random. The new node then forms a link with the randomly chosen node. Repeat the process for newcomers. As more connections are made, it becomes more and more likely that newcomers will pick links going back to the same person. What emerges is a few highly influential people and scores of low-influence people, all connected to one another but certainly not on an equal playing field. The highly connected person gets even more connections because they are already so well-connected. The resulting structure is a scale-free network. How does this relate to shams and indicator properties? There are many ways in which scale-free networks can form – even by random choices – but once formed their relations can take on epistemic properties. Once a mere mortal is elevated to celebrity status, followers begin to look to that person for advice. Recent history is littered with examples of public intellectuals giving opinions outside the bounds of their expertise, which are then adopted by those further out in the network.

To sum up: **BSA** fits intuitions regarding how to evaluate beliefs grounded in the testimony of sham epistemic authorities. Sham authorities co-opt signals of authoritativeness, and there are at least two ways in which signals can fail to reliably identify authorities: through unjust socio-historical circumstances and through high degrees of connectivity in a social network.

4. Authorities Amplifying Implicit Biases: Linguistic Intergroup Bias, Encoding Discriminatory Contents, Conversational Alignment and Audience Attunement Effects

In this section, I review a slice of data from the social psychological literature on the transmission of stereotypes. Specifically, I will introduce four well-established phenomena and suggest how they contribute to the transmission of implicit biases: conversational alignment, Linguistic Intergroup Bias, encoding discriminatory contents, and audience attunement effects. The purpose of this excursion into social psychology is to tell a naturalistic story about social interactions amplifying the effects of on-board implicit biases. That is, the same social relationships do both normative and distal causal explanatory work without inquirers recognizing the causal efficacy of the relationships on their beliefs.²¹

4.1 Mimicking Others: Conversational Alignment

Conversational alignment enables smooth interaction by using the language one's conversational partner has introduced. For example, if I use 'chef' to refer to the person cooking our meal, my conversational partner is likely to use 'chef.' But if I use 'cook,' then my interlocutor will likely use 'cook.' A key element in conversational alignment is the development of a *routine* in conversation (Pickering and Garrod 2004). A routine is an expression with a fixed meaning and the expression occurs more frequently than one would expect given how frequently its constituent words show up in the dialogue (see also Aijmer 1996).²² Examples include common phrases, like 'How are you?' and 'Thanks a lot.' Routines are a dime-a-dozen in conversation. Some estimates hold that 70% of spontaneous speech consists of routines (Altenberg 1993, 1998). Routines don't have to be cemented in a linguistic community. They can be set up 'on the fly' (Pickering and Garrod 2004). For example, a sitting national leader might use 'the previous administration' – as opposed to 'the Jones administration' – to refer to the previous occupants of a network of elected posts. No formal declarations necessary.

Routines are used with their established meaning even though the expressions could be used in non-routine ways (Garrod and Anderson 1987). In one experiment, two subjects were seated on either side of an opaque barrier and each given sheets of paper covered with gridlines. One subject's paper had dots located in various boxes on her grid, and she had to tell the other subject the boxes in which the dots were located. Subjects would refer to the lines and rows of the boxes created by the gridlines to convey information about the dots. Different subjects in different trials would use 'row' and 'line' to refer to horizontal or vertical sequences of boxes. Also, subjects paired 'row' and 'line' with ordinal and non-ordinal terms (e. g. 'first' or 'top'). In a dialogue, subjects used 'row' and 'line' to refer to horizontal or vertical sequences, and then they were consistently used that way. The same goes for pairing 'row' or 'line' with an ordinal or non-ordinal term: once paired for a dialogue, always paired for that dialogue. Subjects would talk about 'the third row' or 'the bottom line.' They would not talk about 'the bottom row' and 'the third line' even though the latter pair of expressions does perfectly well.

4.2 Discriminatory Content Goes Implicit: Linguistic Intergroup Bias and DogWhistles

Not all discriminatory contents need be explicit. Discriminatory content can be implicitly encoded into seemingly innocuous utterances. The first phenomenon describing this is Linguistic Intergroup Bias. Anne Maass et al. (1989, 1995) have found differences in the ways in which groups describe positive and negative behaviors by in-group and out-group members. Positive behaviors by in-group members and negative behaviors by out-groups members tend to be described in terms of abstract character traits, like 'honest' or 'cowardly.' By contrast, negative behaviors by in-group members and positive behaviors by out-group members tend to be described using concrete language, using nouns and verbs to report the event. Crucially, abstract descriptions are perceived as providing more information about the target than concrete descriptions. Abstract descriptions produce the expectation that the target action will be reproduced in the future, which does not necessarily hold for concrete descriptions. And abstract descriptions are more difficult to verify or falsify than concrete ones: compare the evidence needed to validate or refute 'Jamal is aggressive' with 'Jamal hit Susan.' Further, Wigboldus, Semin, and Spears (2000) have found that such linguistic biases also communicate expectations about the behaviors of in-group and out-group members. So if Fred says to Ginger, 'Jamal is aggressive,' then Ginger is more likely to expect aggressive behavior from Jamal and will likely have a harder time disconfirming this view.

The second phenomenon is dogwhistling: deploying words and phrases that resonate with a group of hearers and is missed by outsiders.²³ An example (Saul 2018) comes from George W. Bush's 2003 State of the Union speech:

Yet there's power, wonder-working power, in the goodness and idealism and faith of the American people. (Noah 2004, cited in (Saul 2018))

The dogwhistle is the expression 'wonder-working power' and the group to which it appeals is fundamentalist Christians. The expression refers to the power of Christ, and so Bush's utterance decoded would be:

Yet there's power, the power of Christ, in the goodness and idealism and faith of the American people.

Bush is telegraphing to his fundamentalist constituents that he's one of them. Good for solidarity, and not terribly epistemically problematic. Much more worrisome are what Saul calls 'covert intentional dogwhistles.' Their racist or sexist content has gone undercover. An example is 'inner city.' It functions as a dogwhistle for *Black*. And covert intentional dogwhistles aren't psychologically innocuous. Findings by Horwitz and Peffley (2005) suggest that inserting 'inner city' into the phrase 'violent criminals' affects whether subjects would rather spend money to build prisons or to fund antipoverty programs. The effect was magnified by preexisting racial attitudes. Racial conservatives were much more likely to support building prisons; racial liberals were much more likely to support funding antipoverty programs. Dogwhistles have their effects outside of hearers' conscious awareness. The data suggest that if the discriminatory content of the dogwhistle is made explicit, then the effects dissipate. Dogwhistles, to be efficacious, must bear their contents *implicitly*.

4.3 From Repetition to Belief: Audience Attunement Effects

Linguistic Intergroup Bias and dogwhistles are linguistic mechanisms for encoding biasing information that reinforce existing biases. Conversational alignment has repetition of such biasing phrases. Next is the transition from repeating what others say to believing it. The psychological story is a modern take on an old idea: following others' behaviors changes our own attitudes. Durkheim ([1893]1997) observed this. He saw that rituals create solidarity. He argued that the mechanism connecting solidarity and ritual is shared emotional experience, what he called 'collective effervescence.' Shared emotional experience, in turn, gives rise to a 'collective conscience,' which reinforces

group ties and solidarity. But you don't have to be a sociologist to know that rituals create solidarity and change attitudes. Major world religions often revolve around jointly spoken prayers. Dictators create rituals to bolster solidarity, like Germans of the Third Reich greeting each other in a standardized way. Chicago Cubs fans sing 'Go, Cubs, Go' at every game. The lesson is the same in each case: what we do is a powerful driver of what we believe.

Audience tuning effects are outcomes for speakers of tweaking a message for hearers (cf. Echterhoff et al. 2009). Pitching the message in a specific way affects the speaker as well as the hearers. Higgins and colleagues explore one case of this: the saying-is-believing effect. Asserting that p can lead to believing that p (Tory and Rholes 1978; Tory and McCann 1984; Echterhoff, Higgins, and Levine 2009).²⁴ Higgins and Rholes' classic experiment asks university students to read a story about a man named Donald, whose actions could be described as either reckless or adventurous. Subjects were told that they would be talking about Donald with a partner who either 'kinda liked' or 'kinda disliked' him. Coding of subsequent discussions revealed that participants tended to talk about Donald favorably or not depending on their conversation partners. Subjects with pro-Donald attitudes talked about him in a positive light. And subjects with anti-Donald attitudes talk about him in a negative light. More surprisingly, a follow-up one week later revealed that subjects with pro-Donald partners tended to remember more positive aspects of Donald and endorse more positive claims about him. The same held for subjects with anti-Donald partners. Saying leads to believing.²⁵

The data just reviewed tell this story. Inquirers talk with authorities. The authority introduces a routine reflecting some implicit bias, a routine that the inquirer participates in too. After deploying the routine, inquirers' on-board implicit biases are likely stronger in their influence on belief-formation. But repetitions of this cycle entrench implicit attitudes and their correlate beliefs. It's worth stressing that this isn't the only way in which bias-amplifying information is transmitted. It seems plausible that different kinds of mechanisms are at work in different kinds of interaction. But any psychological story about the mechanisms has to account for how biases-implifying information is passed on unawares.²⁶

5. BSA and Implicitly Biased Beliefs

BSA, I have argued, is an intuitively satisfactory way to evaluate beliefs grounded in the testimony of sham authorities. That is the general case. The specific case is implicitly biased beliefs grounded in the testimony of sham authorities. I have suggested that implicitly biased beliefs get picked up by inquirers through routines and audience-tuning effects. These mechanisms identify one way (of likely many) by which inquirers' implicit biases are reinforced without awareness. Their biases are grounded in the repeated testimony of sham authorities. Consequently, beliefs formed as a result of these implicit biases are well-formed by the inquirer's lights. And on reflection, the inquirer's biased beliefs fit with what they have learned from the authority. But still, the beliefs are epistemically defective. This is exactly what **BSA** says.

What's the takeaway? It is that inquirers can do their level best to listen only to the relevant authorities but still get pulled in by shams. Social psychologist Robert Cialdini admits this about himself:

All my life I've been a patsy. For as long as I can recall, I've been an easy mark for the pitches of peddlers, fundraisers, and operators of one sort or another. True, only some of these people have had dishonorable motives. The others...have had the best of intentions. No matter. With personally disquieting frequency, I have found myself in possession of unwanted magazine subscriptions or tickets to the sanitation workers' ball. (1984, xi)

The appropriate response to Cialdini – or anyone who's pulled in by epistemic con artists – is sympathy. An acknowledgement of 'there, but for the grace of God, go I' in our epistemic lives.

Just the same, when a person forms implicitly biased beliefs after latching onto a sham authority, resulting beliefs are epistemically defective. But we might reasonably believe of them that they're doing their level best to hold true beliefs.

6. Objections and Replies

On the account given in this paper, inquirers latch onto sham authorities because they have been duped by those individuals possessing the relevant indicator properties. Inquirers' beliefs are faulty because they misread the indicator properties or because they picked out unreliable indicator properties, not because they latched onto a bad authority. But on this explanation, the authority-inquirer relationship doesn't do any work.²⁷ The explanation for epistemic fault lies with the individual's failure to assess the evidence properly. The inquirer's belief is epistemically bad because of the inquirer's failure to use reliable indicator properties. Why think they aren't reliable? Because, they are possessed by both genuine and sham authorities. Reliable indicator properties would be possessed only (or only largely) by genuine authorities.

It strikes me that this objection is asking for the impossible: indicator properties that cannot be co-opted. This seems unlikely because being pegged as a genuine epistemic authority carries a range of benefits. If people think I'm an authority in some area, then my opinions are influential. If there's cash involved – like a salary – I'm more likely (though obviously not guaranteed) to get it if I'm believed to be an authority. Glaucon figured this out ages ago: it's enormously beneficial to be *believed* to be authoritative. Given that there's an incentive to be believed by others to be authoritative, the best way to achieve that (short of becoming an actual authority), is to put on the trappings of authority. If you want people to think you're smart, do what the pundits do: interrupt, be mildly condescending, and be intense!²⁸ So if the objection requires coopt-proof indicator properties, then it is asking for more than what us finite knowers are able to do. One might as well ask viceroys not to copy the poisonous-indicating pattern of the monarch.

But now there is another worry in the neighborhood. If indicator properties are co-opted by sham authorities, then there is no way to tell genuine from sham authorities. And if there is no way to tell genuine authorities from shams, then we are doomed.

The pessimism of this corollary worry is unwarranted. Consider the natural phenomenon of mimicry: one species reaping the benefits of another through some kind of copying. For example, the viceroy butterfly is quite tasty for birds. The monarch butterfly? Not so much. Over evolutionary time, viceroys have taken advantage of birds' aversion to monarchs by imitating their coloring. By copying monarchs, viceroys do better. Now this is a case of blending in, and shams want to stand out against non-authorities. So we would expect development of *new* indicator properties to signal differences. Of course this isn't going to happen with butterflies: they do best when not being eaten. But it *does* happen with marketing. According to one (possibly apocryphal) story, Orville Redenbacher increased sales of his popcorn by calling it *gourmet* popping corn. Sticking 'gourmet' in the name enticed people to buy it; his popcorn stood out from the rest. 'Gourmet' acted as an indicator property, indicating a superior product. If 'gourmet' begins to be widely adopted by other popcorn sellers, a new indicator property will need to be found.

The moral is this: the reliability of indicator properties can't *ever* be all or nothing. When some are co-opted, others are found. When the new ones are co-opted, still others will emerge. It seems unreasonable for epistemic agents to stay on the cutting edge of what indicator properties tag genuine authorities because there is no clearly defined cutting edge. Still, the fact that indicators evolve provides reassurance that we have not found ourselves in an epistemic apocalypse. Inquirers get the hang of which properties more frequently pick out genuine authorities.

5. Conclusions

Let us conclude. One reason why implicitly biased beliefs are bad is because they are picked up from sham authorities. The empirical story for how this happens is that inquirers latch onto the sham authorities and then form implicitly biased beliefs. Inquirers pick up on the language used by shams to describe minorities and then come to believe what they're saying.

One swallow doesn't make a summer, nor does one exposure to a sham authority instill implicit biases and attitudes. It takes time to develop biases from interactions with sham authorities. And unfortunately, once the biases are implanted they can be difficult to uproot, particularly if one finds nothing wrong upon reflection. My aim in this paper has been to show inquirers might develop implicit biases through interactions with sham authorities and to suggest that their beliefs are defective, though they themselves find nothing wrong with them. Recognizing that it is reasonable for inquirers to persist in their biased beliefs makes the challenges for egalitarian-minded theorists even more daunting but truer to life.²⁹

Notes

1. Available: <<https://www.theatlantic.com/notes/2017/06/unconscious-bias/529464/>> Retrieved 22 January 2018. Much work has been done on the philosophy and psychology of implicit biases over the last 25 years. A useful collection of philosophical essays is found in Brownstein and Saul (2016). A user-friendly introduction to the science of implicit biases is Banaji and Greenwald (2013). See Dovidio et al. (1997) for an early and influential paper.
2. The starting point for discussion of implicit biases and correlate beliefs is Kahneman's (2009) distinction between 'fast' and 'slow' cognition. Gendler (2008, 2011) argues that implicit biases are 'aliefs' – an automatic and unconscious correlate of beliefs. Mandelbaum (2016) argues that implicit biases are structured propositions, making them more like traditionally-conceived beliefs. Levy (2015) describes them as 'patchy endorsements': mental entities with propositional structure but insufficiently sensitive to other mental representations to count as beliefs. Schwitzgebel (2001) argues that there are in-between beliefs that are best described by dispositional accounts of belief, and some (cf. Egan 2011) suggest that implicit biases might be such in-between beliefs. I will not weigh in on the discussion here, but I presume that there is some difference between an implicit bias and the corresponding beliefs and judgments. Nonetheless, I use 'implicitly biased belief' ('judgment') to indicate that an implicit racial bias is somehow part of the belief's causal history.
3. See Brownstein and Saul (2016) for further developments.
4. The positions developed by Kelly and by Holroyd are subtly different. Those differences, however, are tangential to their shared insight: that cognition threads through the environment.
5. Cf. Madva (2016); and Lassiter and Ballantyne (2017) for similar arguments.
6. Washington and Kelly (2016, 26). Washington, Kelly, and Holroyd all identify developments in extended mind (e.g. Clark and Chalmers 1998) theory as motivating (in part) their views. Thanks to an anonymous reviewer for bringing to my attention that Washington and Kelly are not committed to the metaphysics of extended mind, even if they sound like they are.
7. I focus on the transmission of information that amplifies the epistemic effects of implicit biases. Why? Data suggests that people develop implicit biases during infancy and childhood. For example, Xiao et al. (2018; Liu et al 2015) suggest that babies as young as six months exhibit biases towards same-race faces and biases against different-race faces. So plausibly, when inquirer interactions with authorities result in discriminatory attitudes that are rational upon reflection, it is not that an implicitly racist attitude is transmitted but rather that preexisting implicit attitudes are strengthened. So the mechanisms introduced in this section focus on transmission of information, but it is information that reinforces already-existing implicit biases rather than creating biases where there was none before. Thanks to an anonymous reviewer whose objection prompted this shift from transmission of information for attitude formation to information for bias amplification.
8. The only kind of defect that is excluded from the definition is one in which the agent persists in believing p despite judging on reflection that belief in p is unjustified.
9. Fricker (1994) makes a similar claim.
10. Many thanks to an anonymous reviewer for this suggestion.
11. Readers might notice that **BSA** is closely allied to virtue approaches to knowledge. Greco (1993); Greco (2000) maintains that internalist and externalist dimensions are important for an account of knowledge (cf. Breyer and Greco 2008; Breyer 2013; Sosa 2007, 2009). **BSA** reflects the dual commitment in evaluating beliefs even though it is agnostic on what it takes to elevate the belief to the level of knowledge.

12. Thanks to an anonymous reviewer for highlighting that this view is at home with Zagzebski's exemplarist approach in ethics and epistemology.
13. Hofling et al. (1966). A follow-up study from Rank and Jacobson (1977) failed to replicate Hofling's findings. They suspect this was because nurses were allowed to confer with one another before delivering the medicine, the medicine itself was different (and better known), and nurses were more willing to challenge hospital authority.
14. How is **Con Artist** different from the creationist teacher of Lackey (1999)? Ignore that the creationist teacher conveys truths and the con artist some falsehoods. Lackey's case is one in which the transmitting agent really does have some sort of epistemic standing. By hypothesis she's a teacher. Our con artist, by hypothesis, is not. He's only faking it, even if convincingly. So even if you're inclined to agree with Lackey that hearers can acquire knowledge from agents lacking that knowledge, **Con Artist** is not such a case. This is true even though neither the teacher nor the con artist believes the information being transmitted. Thanks to Chris Tucker for this concern.
15. Thanks to an anonymous reviewer for suggesting I make these points about Mac as a sham more explicit.
16. See also Goldman (Goldman 2009; forthcoming). These sources are: (1) arguments supporting their own views and critiquing their rivals, (2) agreement from other experts, (3) appraisal by 'meta-experts,' (4) evidence of interests and biases, and (5) experts' track-records. For discussion and critique of Goldman's position, see Matheson (2005); Coady (2006); Scholz (2009); and Zagzebski (2012).
17. Thanks to Nathan Ballantyne for suggesting this example.
18. For example, Anderson, Green, and Payne (2009) note that doctors tend to undertreat Blacks (relative to Whites) when reporting pain. A natural way to read this is that Whites are authorities with respect to their own pain experiences while Blacks are not.
19. See Zollman (2007, 2012) for detailed discussions of epistemic networks.
20. Cf. Barabási and Albert (1999). 'Scale-free' because as you zoom in on the larger network (i.e. adjust the scale) the structure stays more or less the same.
21. While this section is speculative, similar conclusions have been reached independently by Beukeboom (2014), whose work was made known to the author only after this paper had been written and was under review.
22. This rules out an expression E becoming a routine in conversations about E. 'Ludwig Wittgenstein' isn't a routine in a conversation about Wittgenstein since the whole expression likely shows up as often as the constituent elements show up.
23. It's worth noting that routines can take advantage of ordinarily innocuous words to accomplish this. For example, one might use the expression, 'people like that' or 'those people' to refer to Blacks or Muslims (Edwards 2004).
24. The speaker can't be bribed, coerced, or otherwise compelled into speaking (Klass 1978). It seems plausible that there are other restrictions: talking about genocide in a positive light isn't (I think) going to make me think favorably of it. But these are empirical questions for which more work needs to be done.
25. Which is why I now require all my students to tell me how much they like the class just before they complete their course evaluations.
26. The ways in which agents influence the judgments and beliefs of other agents are legion. Related findings include source credibility (Hovland and Weiss 1951; Chaiken and Maheswaran 1994), attractiveness of the sham authority (Chaiken 1979), and judgments of trustworthiness (Levin, Whitener, and Cross 2006).
27. Thanks to Hilary Kornblith for this objection.
28. Eric Schwitzgebel has a fine discussion on 'being good at seeming smart' at his blog *The Splintered Mind*: <http://schwitzsplinters.blogspot.com/2010/03/on-being-good-at-seeming-smart.html>.
29. Thanks to audiences at the 2017 Eastern Division APA and the 2016 NYC Epistemology/Psychology conference at Fordham University. Thanks also to anonymous reviewers at this journal for many helpful suggestions.
30. Searching trends for the relative popularity of 'unconscious bias,' 'unconscious racism,' 'implicit bias test,' and 'unconscious prejudice' shows that 'implicit bias' is more searched for relative to these other expressions. Thanks to Peter Kirwin for the suggestion.
31. There are tools available for getting and analyzing search volume. These cost money; Google Trends and census data don't.
32. Thanks to an anonymous reviewer for pressing clarification of the methodology.
33. This includes the time period of most intense searches for 'implicit bias' but excludes a surge of interest in 'powerball' in late November 2016.
34. There are limitations to the data, of course. This only looks at total searches and there might be people googling 'powerball' more than once. And we only have what people are googling – that's to say what people are interested in finding out. So it might well be that there are more people who know about implicit biases than the data suggests. This data isn't available, so Google Trends has to be our proxy. But given the relative popularity of searches for 'implicit bias,' optimism about widespread awareness of or interest in implicit biases is misplaced.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Charles Lassiter is an assistant professor of philosophy at Gonzaga University. His primary areas of research are metaphysics of mind and social epistemology. He draws on empirical findings to develop a neo-Aristotelian view of mind and knowledge.

References

- Adler, J. 1997. "Lying, Deceiving, or Falsely Implicating." *The Journal Of Philosophy* 94 (9): 435–452.
- Aijmer, K. 1996. *Conversational Routines in English: Convention and Creativity*. New York, NY: Routledge.
- Altenberg, B. "Recurrent Word Combinations in Spoken English." *Proceedings of the Fifth Nordic Association for English Studies Conference*. Reykjavik: University of Iceland, 1993.
- Altenberg, B. 1998. "On the Phraseology of English: The Evidence of Recurrent Word-Combinations." In *Phraseology*, edited by A. Cowie, 101–122. New York, NY: Oxford University Press.
- Anderson, K. O., C. R. Green, and R. Payne. 2009. "Racial and Ethnic Disparities in Pain: Causes and Consequences of Unequal Care." *The Journal of Pain* 10 (12): 1187–1204. doi:10.1016/j.jpain.2009.10.002.
- Ballantyne, N. "Epistemic Trespassing," *Mind* (forthcoming)
- Banaji, M., and A. Greenwald. 2013. *Blindspot: Hidden Biases of Good People*. New York: Bantam Books.
- Barabási, A.-L., and R. Albert. 1999. "Emergence of Scaling in Random Networks." *Science* (New York, NY) 286: 509–512.
- Beukeboom, C. J. 2014. "Mechanisms of Linguistic Bias: How Words Reflect and Maintain Stereotypic Expectancies (Chapt.)." In *Social Cognition and Communication*, edited by J. Laszlo, J. Forgas, and O. Vincze, 313–330. New York, NY: Psychology Press.
- Breyer, D. 2013. "Knowledge, Credit, and Cognitive Agency." *Pacific Philosophical Quarterly* 94 (4): 503–528. doi:10.1111/papq.2013.94.issue-4.
- Breyer, D., and J. Greco. 2008. "Cognitive Integration and the Ownership of Belief: Response to Bernecker." *Philosophy and Phenomenological Research* 76 (1): 173–184. doi:10.1111/phpr.2008.76.issue-1.
- Brownstein, M. and J Saul, eds., 2016. *Implicit Bias and Philosophy, Volumes 1 & 2*. New York, NY: Oxford University Press.
- Chaiken, S. 1979. "Communicator Physical Attractiveness and Persuasion." *Journal of Personality and Social Psychology* 37 (8): 1387–1397. doi:10.1037/0022-3514.37.8.1387.
- Chaiken, S., and D. Maheswaran. 1994. "Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment." *Journal of Personality and Social Psychology* 66 (3): 460–473. doi:10.1037/0022-3514.66.3.460.
- Cialdini, R. 1984. *Influence: The Psychology of Persuasion*. New York, NY: HarperCollins Publishing.
- Clark, A., and D. Chalmers. 1998. "The Extended Mind." *Analysis* 58: 7–19. doi:10.1093/analys/58.1.7.
- Coady, D. 2006. "When Experts Disagree." *Episteme: A Journal of Social Epistemology* 3: 68–79. doi:10.3366/epi.2006.3.1-2.68.
- DePaul, M. R. 1993. *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry*. New York, NY: Routledge.
- Dovidio, J. F., K. Kawakami, C. Johnson, B. Johnson, and A. Howard. 1997. "On the Nature of Prejudice: Automatic and Controlled Processes." *Journal of Experimental Social Psychology* 33: 510–540. doi:10.1006/jesp.1997.1331.
- Durkheim, E. [1893]1997. *The Division of Labor in Society*. Trans. W.D. Hall. New York: Free Press.
- Echterhoff, G., E. T. Higgins, and J. M. Levine. 2009. "Shared Reality: Experiencing Commonality with Others' Inner States about the World." *Perspectives on Psychological Science* 4: 496–521. doi:10.1111/j.1745-6924.2009.01161.x.
- Echterhoff, Gerald, S. Lang, N. Krämer, and E. T. Higgins 2009. "Audience-Tuning Effects on Memory: The Role of Audience Status in Sharing Reality". *Social Psychology* 40: 150–163. doi: 10.1027/1864-9335.40.3.150.
- Edwards, D. 2004. "Analyzing Racial Discourse: The Discursive Psychology of Mind–World Relationships." In *Analyzing Race Talk: Multidisciplinary Perspectives on the Research Interview*, edited by H. van den Berg, M. Wetherell, and H. Houtkoop-Steenstra, 31–48. New York, NY: Cambridge University Press.
- Egan, A. 2011. "Comments on Gendler's 'The Epistemic Costs of Implicit Bias,'" *Philosophical Studies* 156: 65–79. doi:10.1007/s11098-011-9803-5.
- Fricke, E. 1994. "Against Gullibility." In *Knowing from Words*, edited by B. K. Matilal and A. Chakrabarti, 125–161. Dordrecht: Kluwer Academic Publishers.
- Fricke, M. 1998. "Rational Authority and Social Power: Towards a Truly Social Epistemology." *Proceedings of the Aristotelian Society* 2: 159–177. doi:10.1111/1467-9264.00030.

- Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. New York, NY: Oxford University Press.
- Garrod, S., and A. Anderson. 1987. "Saying What You Mean in Dialogue: A Study in Conceptual and Semantic Co-Ordination." *Cognition* 27 (2) : 181–218. doi:10.1016/0010-0277(87)90018-7.
- Gendler, T. 2008. "Alief and Belief." *Journal of Philosophy* 105 (10): 633–663. doi:10.5840/jphil20081051025.
- Gendler, T. 2011. "On the Epistemic Costs of Implicit Bias." *Philosophical Studies* 156 (1): 33–63. doi:10.1007/s11098-011-9801-7.
- Goldman, A. 2001. "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63: 85–110. doi:10.1111/phpr.2001.63.issue-1.
- Goldman, A. 2009. "Social Epistemology: Theory and Applications." *Royal Institute of Philosophy Supplement* 64: 1–18. doi:10.1017/S1358246109000022.
- Goldman, A. forthcoming. "Expertise." *Topoi*. doi:10.1007/s11245-016-9410-3.
- Greco, J. 1993. "Virtues and Vices of Virtue Epistemology." *Canadian Journal of Philosophy* 23 (3): 413–432. doi:10.1080/00455091.1993.10717329.
- Greco, J. 2000. *Putting Skeptics in Their Place: The Nature of Skeptical Arguments and Their Role in Philosophical Inquiry*. New York, NY: Cambridge University Press.
- Hofling, C. K., E. Brotzman, S. Dalrymple, N. Graves, and C. Bierce. 1966. "An Experimental Study of Nurse-Physician Relations." *Journal of Nervous and Mental Disease* 143: 171–180.
- Holroyd, J. 2012. "Responsibility for Implicit Bias." *Journal of Social Philosophy* 43 (3): 274–306. doi:10.1111/j.1467-9833.2012.01565.x.
- Holroyd, J. 2015. "Implicit Bias, Awareness, and Imperfect Cognitions." *Consciousness and Cognition* 33: 511–523. doi:10.1016/j.concog.2014.08.024.
- Holroyd, J., and D. Kelly. 2016. "Implicit Bias, Character, and Control." In *From Personality to Virtue: Essays on the Philosophy of Character*, edited by A. Masala and J. Webber, 106–133. New York: Oxford University Press.
- Holroyd, J., R. Scaife, and T. Stafford. 2017. "Responsibility for Implicit Bias." *Philosophy Compass* 12. doi:10.1111/phc3.12410.
- Horowitz, J., and M. Peffley. 2005. "Playing the Race Card in the Post-Willie Horton Era: The Impact of Racialized Code Words on Support for Punitive Prison Policy." *The Public Opinion Quarterly* 69: 99–112. doi:10.1093/poq/nfi004.
- Hovland, C. I., and W. Weiss. 1951. "The Influence of Source Credibility on Communication Effectiveness." *Public Opinion Quarterly* 15 (4): 635–650. doi:10.1086/266350.
- Kahneman, D. 2009. *Thinking, Fast and Slow*. New York: Farrar, Strauss, and Giroux.
- Klass, E. T. 1978. "The Psychological Effects of Immoral Actions: The Experimental Evidence." *Psychological Bulletin* 85: 756–771. doi:10.1037/0033-2909.85.4.756.
- Lackey, J. 1999. "Testimonial Knowledge and Transmission." *Philosophical Quarterly* 50 (197): 471–490. doi:10.1111/1467-9213.00154.
- Lai, C. K., K. M. Hoffman, and B. A. Nosek. 2013. "Reducing Implicit Prejudice." *Social and Personality Psychology Compass* 7 (5): 315–330. doi:10.1111/spc3.12023.
- Lassiter, C., and N. Ballantyne. 2017. "Implicit Racial Bias and Epistemic Pessimism." *Philosophical Psychology* 30: 79–101. doi:10.1080/09515089.2016.1265103.
- Levin, D. Z., E. M. Whitener, and R. Cross. 2006. "Perceived Trustworthiness of Knowledge Sources: The Moderating Impact of Relationship Length." *Journal of Applied Psychology* 91 (5): 1163–1171. doi:10.1037/0021-9010.91.5.1163.
- Levy, N. 2015. "Neither Fish nor Fowl: Implicit Attitudes as Patchy Endorsements." *Nous* (Detroit, Mich.) 49: 800–823. doi:10.1111/nous.12074.
- Levy, N. 2017. "Implicit Bias and Moral Responsibility: Probing the Data." *Philosophy and Phenomenological Research* 47: 3–26.
- Liu, S., N. G. Xiao, P. C. Quinn, D. Zhu, G. Liezhong, O. Pascalis, and K. Lee. 2015. "Asian Infants Show Preference for Own-Race but Not Other-Race Female Faces: The Role of Infant Caregiving Arrangements." *Frontiers in Psychology* 6. <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00593/full>.
- Maass, A., A. Milesi, S. Zabbini, and D. Stahlberg. 1995. "Linguistic Intergroup Bias: Differential Expectancies or In-Group Protection?" *Journal of Personality and Social Psychology* 68 (1): 116–126.
- Maass, A., D. Salvi, L. Arcuri, and G. R. Semin. 1989. "Language Use in Intergroup Contexts: The Linguistic Intergroup Bias." *Journal of Personality and Social Psychology* 57 (6): 981–993.
- Madva, A. 2016. "A Plea for Anti-Anti-Individualism: How Oversimple Psychology Misleads Social Policy." *Ergo* 3: 701–728.
- Mandelbaum, E. 2016. "Attitude, Inference, Association: On the Propositional Structure of Implicit Bias." *Nous* 50 (3): 629–658. doi:10.1111/nous.2016.50.issue-3.
- Mast, M. S. 2002. "Dominance as Expressed and Inferred Through Speaking Time." *Human Communication Research* 28 (3): 420–450. doi:10.1111/hcre.2002.28.issue-3.
- Matheson, D. 2005. "Conflicting Experts and Dialectical Performance: Adjudication Heuristics for the Layperson." *Argumentation* 19: 145–158. doi:10.1007/s10503-004-7041-0.
- Medina, J. 2010. "The Relevance of Credibility Excess in a Propositional View of Epistemic Injustice: Differential Epistemic Authority and the Social Imaginary." *Social Epistemology* 25: 15–35. doi:10.1080/02691728.2010.534568.

- Mugg, J. 2013. "What are the Cognitive Costs of Racism? A Reply to Gendler." *Philosophical Studies* 166 (2): 217–229. doi:10.1007/s11098-012-0036-z.
- Noah, T. "Why Bush Opposes Dred Scott," (2004): http://www.slate.com/articles/news_and_politics/chatbox/2004/10/why_bush_opposes_dred_scott.html
- Pickering, M. J., and S. Garrod. 2004. "Toward a Mechanistic Psychology of Dialogue." *Behavioral and Brain Sciences* 27 (2): 169–225.
- Rank, S. G., and C. K. Jacobson. 1977. "Hospital Nurses' Compliance with Medication Overdose Orders: A Failure to Replicate." *Journal of Health and Social Behavior* 18: 188–193.
- Rogoff, B. 1990. *Apprenticeship in Thinking*. New York, NY: Oxford University Press.
- Saul, J. 2013. "Implicit Bias, Stereotype Threat and Women in Philosophy." In *Women in Philosophy: What Needs to Change?*, edited by F. and K Hutchinson, 39–60. New York NY: Oxford University Press.
- Saul, J. 2018. "Dogwhistles, Political Manipulation, and Philosophy of Language." In *New Works on Speech Acts*, edited by D. Fogal, M. Cross, and D. Harris, 360–383. New York, NY: Oxford University Press.
- Scholz, O. 2009. "Experts: What They are and How We Recognize Them—A Discussion of Alvin Goldman's Views." *Grazer Philosophische Studien* 79: 187–205. doi:10.1163/18756735-90000864.
- Schwitzgebel, E. 2001. "In-between Believing." *The Philosophical Quarterly* 51: 76–82. doi:10.1111/phiq.2001.51.issue-202.
- Shapin, S. 1994. *A Social History of Truth — Civility and Science in Seventeenth-Century England*. Chicago: University of Chicago Press.
- Sosa, E. 2007. *Apt Belief and Reflective Knowledge, Volume 1: A Virtue Epistemology*. Oxford: Oxford University Press.
- Sosa, E. 2009. *Apt Belief and Reflective Knowledge, Volume II: Reflective Knowledge*. Oxford: Oxford University Press.
- Steinbeck, J. 1936. In *Dubious Battle*. New York, NY: Random House.
- Stephens-Davidowitz, S. 2017. *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us about Who We Really Are*. New York, NY: HarperCollins.
- Tory, H. E., and C. D. McCann. 1984. "Social Encoding and Subsequent Attitudes, Impressions, and Memory: "Context-Driven" and Motivational Aspects of Processing." *Journal of Personality and Social Psychology* 47 (1): 26–39. doi:10.1037/0022-3514.47.1.26.
- Tory, H. E., and W. S. Rholes. 1978. "Saying Is Believing": Effects of Message Modification on Memory and Liking for the Person Described." *Journal of Experimental Social Psychology* 14 (4): 363–378. doi:10.1016/0022-1031(78)90032-x.
- Washington, N., and D. Kelly. 2016. "Who's Responsible for This? Moral Responsibility, Externalism, and Knowledge about Implicit Bias." In *Implicit Bias and Philosophy II: Moral Responsibility, Structural Injustice, and Ethics*, edited by M. Brownstein and J. Saul, 12–36. New York, NY: Oxford University Press.
- Wigboldus, D. H. J., G. R. Semin, and R. Spears. 2000. "How Do We Communicate Stereotypes? Linguistic Bases and Inferential Consequences." *Journal of Personality and Social Psychology* 78 (1): 5–18.
- Xiao, N. G., P. Quinn, Y. Shaoling, G. Liezhong, O. Pascalis, and K. Lee. 2018. "Older but Not Younger Infants Associate Own-Race Faces with Happy Music and Other-Race Faces with Sad Music." *Developmental Science* 21.2. doi:10.1111/desc.12537.
- Zagzebski, L. 2012. *Epistemic Authority: A Theory of Trust, Autonomy, and Authority, in Belief*. New York, NY: Oxford University Press.
- Zagzebski, L. 2017. *Exemplarist Moral Theory*. New York, NY: Oxford University Press.
- Zollman, K. 2007. "The Communication Structure of Epistemic Communities." *Philosophy of Science* 74: 574–587. doi:10.1086/525605.
- Zollman, K. 2012. "Network Epistemology: Communication in Epistemic Communities." *Philosophy Compass* 8: 15–27. doi:10.1111/j.1747-9991.2012.00534.x.

Appendix 1

I strongly suspect most Americans and western Europeans are not familiar with the concept of implicit bias. Social scientists have yet to conduct a survey about how widespread knowledge about implicit biases is. Despite lacking this data, we can get a sense of how widespread knowledge about implicit biases is by looking at the relative volume of Google searches over time. Google searches as a source of data about mental states and attitudes is discussed in Stephens-Davidowitz (2017).

For an expression E and a period of time t , a Google Trends search for E during t finds the unit of time (day, week, hour) in which searches for E were highest relative to all searches for E in t . This gets a score of 100. Popularity of E during all other time units in t is then quantified relative to this high water mark. Look at Figure 1.³⁰

Peak activity during this five-year span occurred towards November 2016. That is to say: from January 2013 to December 2018, the great number of searches for 'implicit bias' happened in fall 2016. All popularity of other search dates are ranked relative to this one. In case you are curious why fall 2016 was the peak time, it is most likely due the US presidential debate in late September 2016 when Hillary Clinton mentioned implicit biases in response to a question put to her by Donald Trump.

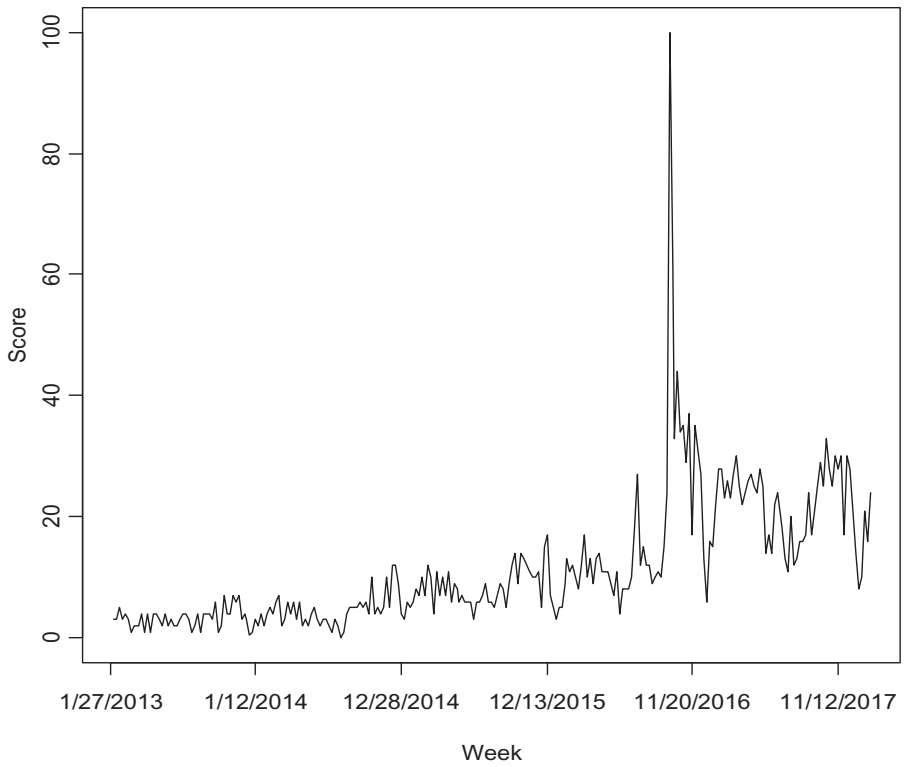


Figure 1. Relative popularity of 'implicit bias' from 2013 to 2018.

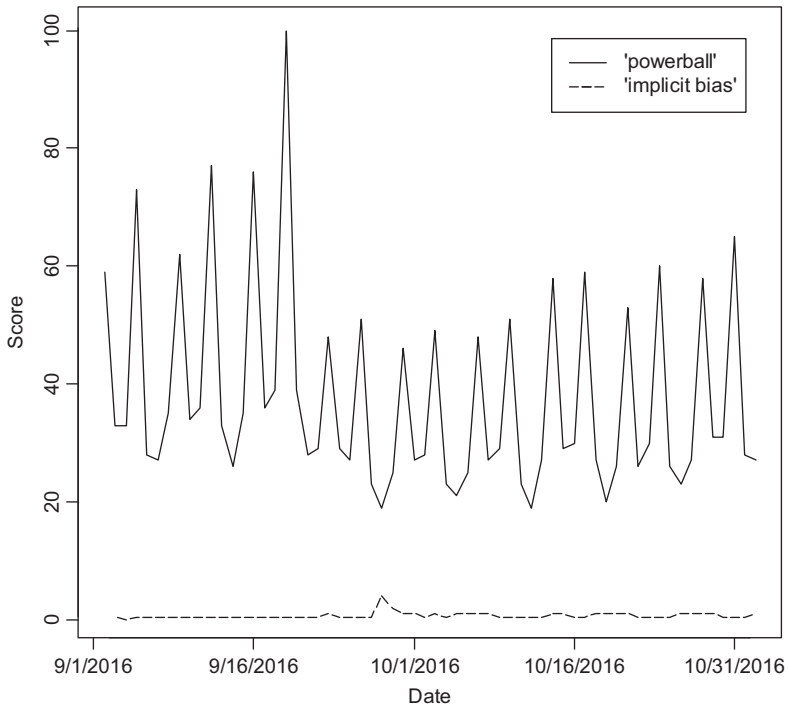


Figure 2. 'Powerball' vs 'implicit bias.'

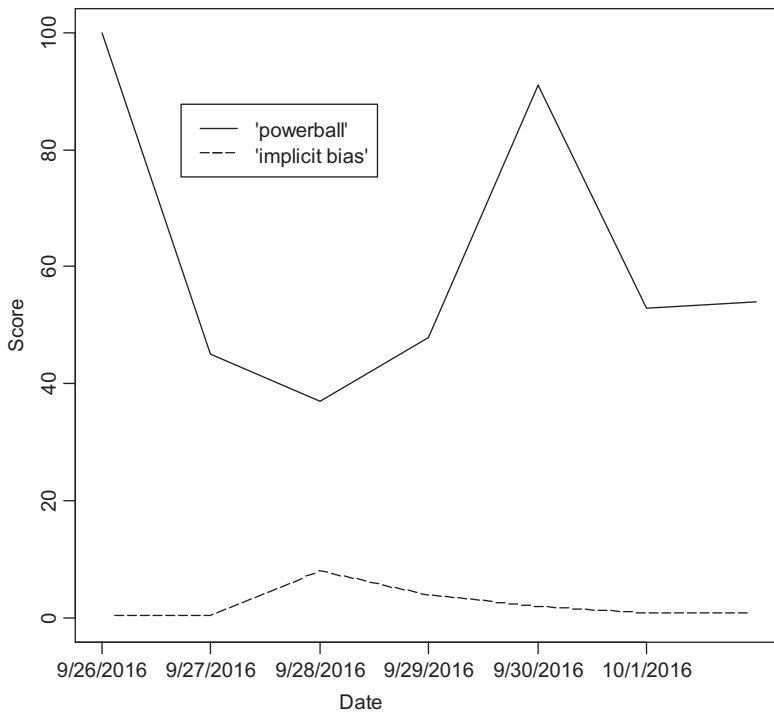


Figure 3. 'Powerball' vs 'implicit bias.'

These data alone do not give us much information about the intensity of searches for 'implicit bias' relative to other terms. All we know is that the greatest number of people searched for 'implicit bias' in fall 2016. Maybe the number of searches was in the hundreds for that time. Maybe it was in the hundreds of millions. With the Trends results for 'implicit bias' from 2013–2018, there's no way to tell.³¹ Trends delivers relativized volume, not absolute volume of, of searches. We can sharpen the picture by comparing Trend results for 'implicit bias' against the most-searched-for term of 2016: 'powerball' – a lottery game in the US where winners get hundreds of millions of dollars. How does this sharpen the picture? While we still don't have an absolute value for searches, it tells us how popular 'implicit bias' is compared to the most popular term. Given that there are, in 2018, roughly 325 million citizens in the US, the proportion of searches for 'implicit bias' to 'powerball' provides a maximum value for absolute search volume for 'implicit bias.' Suppose that the ratio of searches for 'implicit bias' to searches for 'powerball' is 1:100 and assume that half the US population searched for 'powerball' and the ratio is 1:100 – the range of absolute search volume would be in the ballpark of 0.81 to 3.25 million searches.³²

Figure 2 shows compares the popularity of searches for 'powerball' and 'implicit bias' covering the period during the surge of interest: from 1 September 2016 to 1 November 2016.³³ Notice that 'implicit bias' hardly registers. The surge from fall 2016 in Figure 1 is the small bump in Figure 2. Now let us zoom in to the week where 'implicit bias' was at its peak popularity: 25 September to 1 October 2016. This is shown in Figure 3. Notice that 'powerball' still tops 'implicit bias' for popularity.

So why all the big data?³⁴ My suggestion is this. Lacking data about how widespread the concept *implicit bias* is, our next best move is to look at what people are looking up: after all, people can't google something if they don't first have the expression to google. The Google Trends data acts as a rough proxy to tell us how popular a term is relative to others, and the data show that the number of people interested in knowing if they've won the lottery vastly outnumbers the number of people interested in implicit biases. That's to say, curiosity about and awareness of implicit biases are relegated to a small fraction of people in the US.