

Erin Beeghly

Please do not cite without permission from author.

Comments welcome: [erin.beeghly@gmail.com](mailto:erin.beeghly@gmail.com)

Forthcoming in *Introduction to Implicit Bias: Knowledge, Justice, and the Social Mind*

## Epistemic Appraisals of Implicit Bias: Two Metaphors

**Chapter Overview:** If you care about securing knowledge, what is wrong with being biased? Often it is said that we are less accurate and reliable knowers due to implicit biases. Likewise, many people think that biases reflect inaccurate claims about groups, are based on limited experience, and are insensitive to evidence. Chapter 3 investigates objections such as these with the help of two popular metaphors: bias as fog and bias as shortcut. Guiding readers through these metaphors, I argue that they clarify the range of knowledge-related objections to implicit bias. They also suggest that there will be no unifying problem with bias from the perspective of knowledge. That is, they tell us that implicit biases can be wrong in different ways for different reasons. Finally, and perhaps most importantly, the metaphors reveal a deep—though perhaps not intractable—disagreement among theorists about whether implicit biases can be good in some cases when it comes to knowledge.

### 1. Introduction

In the fall of 2016, *The New York Times* published a six-part series of videos—*Who Me? Biased?*—about implicit bias and race. Part of the challenge of these videos was to convey, as quickly and effectively as possible, what implicit bias is and why anyone should care about it.

To meet this challenge, filmmaker Saleem Reshamwala used metaphor. In the first video, he explained to viewers that biases are “little mental shortcuts that hold judgments that you might not agree with” (Reshamwala 2016). One of his guests, psychologist Dolly Chugh, likened implicit bias to a “fog that you’ve been breathing in your whole life.”

These two metaphors—bias as fog and bias as shortcut—are two of many metaphors that one finds in popular and academic writing about implicit bias. One TedX presenter explains to her audience that implicitly biased people live “in the matrix”—a reference to a 1990s film in which characters believe that they are in touch with reality but their experiences are in fact generated by a computer (Funchess 2014). In *Blindspot: Hidden Biases of Good People*, psychologists Mahzarin Banaji and Anthony Greenwald write that implicit biases are “mind bugs,” deploying a computer programming metaphor (Banaji & Greenwald 2013: 13). Nilanjana Dasgupta uses the image of a mirror, writing that biases are “mirror-like reflections” of the social world (Dasgupta 2013: 240). Philosophers get in on the action too. In the edited volume *Implicit Bias & Philosophy*, Louise Anthony says that implicit biases are “mental machinery” used to reliably track “stable but contingent features of our environment” (Antony 2016: 181).

In this essay, I examine two of the most striking metaphors mentioned above: bias as fog and bias as shortcut. I argue that each metaphor makes a distinctive claim about the relationship between bias, knowledge, and error. They also clarify the range of knowledge-related objections to implicit bias. That is, the metaphors tell us that implicitly biased judgments can be wrong in different ways for different reasons. Likewise, the metaphors reveal a deep disagreement among theorists of bias. According to some theorists, implicitly biased judgments are always bad from

the perspective of knowledge. According to others, biased judgments can be rational in certain cases and may even help us to gain knowledge about the world.

## 2. What is an Epistemic Objection?

The word “epistemic” comes from the Greek word “epistēmē,” meaning “knowledge.” Epistemic objections are objections concerning knowledge (cf. Introduction).

Imagine a teenager and a parent having an early morning conversation. The parent says to the teen, “You should get ready to go. The bus will be here at 8:30am.” The teen replies, “That’s not true. The bus is coming at 8:50.” The teen is making an epistemic objection. She is arguing that her father’s belief is *false*. Suppose the parent tries to defend his claim by saying, “I know the bus schedule. Get ready.” The teen replies: “You shouldn’t trust your memory. I just checked online for the latest bus times. The next bus is coming at 8:50.” These are epistemic objections too. The teen asserts that her father’s belief is *unwarranted* by the evidence; moreover, she points out that the way in which he formed the belief is *unreliable* or perhaps *less reliable* than the way in which she formed hers. Suppose the conversation turns even more heated. The father commands the teen to leave the house immediately, lest she be late. The teen yells: “You are being irrational. The bus is not coming for another twenty minutes.” This is epistemic criticism in its strongest form. The teen is asserting that the father is holding onto a false belief despite having compelling evidence that he should give it up.

As this example suggests, epistemic criticism is a constant feature of human life. Humans constantly evaluate each other in epistemic terms, and we are capable of reflecting on the ways in which our own judgments and processes of reasoning could be improved.

### 3. Metaphors & The Epistemic Significance of Bias

What are the best epistemic objections to implicit bias? Is there a single objection that always applies when people make biased judgments? Might there be cases in which implicitly biased judgments are permissible or even good from the perspective of knowledge? Or, are biased judgments necessarily bad from an epistemic point of view?

When trying to understand how one might answer these questions, it is useful to start with metaphors. Metaphors often serve as what philosopher Elisabeth Camp calls *interpretative frames*. Camp explains:

...a representation [which could be visual or linguistic in nature] functions as a frame when an agent uses it to organize their overall intuitive thinking...a frame functions as an overarching, open-ended interpretative principle: it purports to determine for any property that might be ascribed to the subject, both *whether* and *how* it matters (Camp forthcoming: 5).

Two features of frames are especially important. First, they make certain features of a person or thing *salient* in cognition or perception. Also, Camp says, metaphors attribute *centrality* to

certain features of a person, group, or thing. For example, they identify some features of a thing as having special causal powers and as especially important to making the thing what it is (6).

To better understand these two effects of metaphor, consider an example. In the play *Romeo and Juliet*, Romeo says about his love interest, “Juliet is the sun.” Romeo’s use of metaphor renders a specific feature of Juliet *salient*: her stunning physical beauty, i.e., her “hotness.” As Juliet’s beauty becomes salient, other features of her recede into the background. The metaphor also attributes *centrality* to Juliet’s beauty. Her desire-inducing physical appearance is what makes Juliet worthy of Romeo’s love and devotion. It is a driver of drama in the play and is asserted to be crucially important to making Juliet the special person she is.

Here is what Camp’s view suggests. The metaphors used to talk about implicit bias are not mere rhetorical flourishes whose main purpose is to make discussions of implicit bias more exciting or accessible. On her view, metaphors are cognitively crucial. They reveal how speakers intuitively conceptualize a phenomenon like implicit bias. Camp puts the point like this: metaphors—and interpretative frames more generally—provide the “‘intuitive ‘mental setting’ (Woodfield 1991, 551) or background against which specific beliefs and questions are formulated” (3; Lakoff & Johnson 1980).

If she is correct, investigating the metaphors associated with implicit bias will tell us something interesting about how theorists intuitively understand bias and its epistemic significance. These metaphors will also give us a vivid entry point into thinking about when and why implicit biased judgments are problematic, from an epistemic point of view.

#### 4. Living in a Fog

Start with the metaphor of fog. Fog is “a state of the weather in which thick clouds of water vapor or ice crystals suspended in the atmosphere form at or near the earth's surface, obscuring or restricting visibility to a greater extent than mist” (OED 2017a). At the website for *Take the Lead*—an organization that promotes women in business—writer Michele Weldon says:

Implicit gender bias has hung around women leaders in the workplace in nearly every imaginable sector and discipline for generations. The bias surrounds the workplace culture in a fog at times thick and impenetrable, and at other times, a mist that only feels instinctively palpable (Weldon 2016).

If implicit bias is fog, the effect is obvious: people will have a hard time perceiving the world as it really is because of it (cf Siegel, Ch. 5). Sensory perceptions of other people and the world become fuzzy, impaired. If you look at someone through a fog, for example, you might think, “I can’t really see you as you are. I see the fog, and I see a blurry version of you.”

In the *Times* video, the metaphor is taken a step further. Not only does fog obstruct visual and auditory perception, it becomes internalized. “We’ve all grown up in a culture,” says Chugh, “with media images, news images, conversations we’ve heard at home, and education...think of that as a fog that we’ve been breathing our whole lives, we never realized what we’ve been taking in.” That fog, Reshamwala adds, “causes associations that lead to biases.” For example, when you hear *peanut butter*, you think *jelly*. That association exists because peanut butter and

jelly are typically paired together in our culture. Similarly, Chugh observes, “in many forms of media, there is an overrepresentation of black men and violent crime being paired together.” The result is, as educational scholar Shaun Harper puts it in the video, “deep down inside we have been taught that black men are violent and aggressive and not to be trusted, that they’re criminals, that they’re thugs.”

Remember that metaphors, like all interpretative frames, are supposed to do two important things. First, they render certain aspects of a phenomenon more *salient* in cognition, and they assert claims about the causal *centrality* of certain properties.

What becomes salient if we think of implicit bias as fog? Here is one thing: its epistemic badness. Bias, if it is a kind of fog, clouds vision and distorts hearing. Cognitive fog is no better. When people talk about the fog of war, what they mean is that war creates an environment where soldiers cannot think clearly, cannot accurately evaluate risks, and cannot make good decisions. Oppressive social conditions create something similar, according to this metaphor: *the fog of oppression*.

This way of thinking about bias resonates, in particular, with theorists of oppression, especially philosophers of race. In *The Racial Contract*, for example, Charles Mills describes conditions of white supremacy as requiring “a certain schedule of structured blindness and opacities in order to establish and maintain the white polity” (Mills 1999: 19). Applied to implicit bias, the thought is something like this. Many folks today don’t explicitly endorse racist, sexist, classist, or otherwise prejudiced views. They reject them. However, they—especially but not exclusively members of

dominant groups—absorb these problematic views and, as a result, think and act in ways that reproduce conditions of injustice. Yet they do not recognize themselves as being part of the problem; sometimes they don't even realize that there is a problem. Mills prefers to use the metaphor of *collective hallucination* to describe this state of ignorance (18). But fog is supposed to function similarly. For oppressive conditions to persist, the fog/hallucination must persist.

How does bias—understood as fog—frustrate accurate vision and cognition? Here is one possibility: through group stereotypes. Stereotypes fit the description of “fog” at least in one way. They exist in the world and not just in individuals' minds. During the 18<sup>th</sup> and 19<sup>th</sup> centuries, the word “stereotype” was a technical term in the printing industry. Stereotypes were the metal plates used in printing presses (*see Figure A*). The process of creating these plates was called “stereotyping.” The first book “stereotyped” in the U.S. was the New Testament in 1814. By the early 20<sup>th</sup> century, every newspaper office had a stereotyping room, where both full-page plates for regular pages and smaller plates for advertising were produced. Common images and phrases reproduced by this technology, i.e., by stereotyping were also deemed “stereotypical.” In the Figure B, for instance, one sees a stereotypical Japanese person as represented in U.S. World War II propaganda. Remember what Chugh says: media and news images partially constitute the fog. This claim dovetails with assertions by feminist scholars that stereotypes exist “in the social imaginary” (Fricker 2007) and in “the mind of the world” (Siegel 2017).

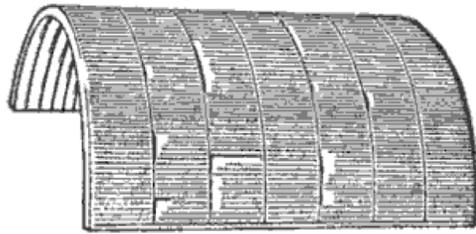


FIG. 5  
Curved Stereotype for Newspaper Page

Figure A: Example of Stereotype Plate for Printing Press

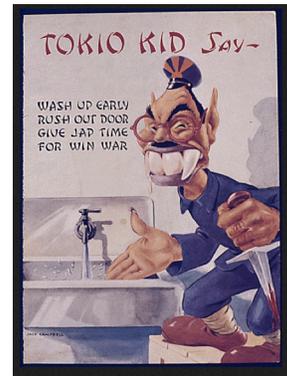


Figure B: Stereotypical Image of Japanese Person in WWII Propaganda

A very common claim about stereotypes—which would explain why they constitute a kind of fog—is that stereotypes are necessarily false or misleading. As philosopher Lawrence Blum notes,

By and large, the literature on stereotypes (both social psychological and cultural) agrees that the generalizations in question are false or misleading, and I think this view generally accords with popular usage...The falseness of a stereotype is part of, and is a necessary condition of, what is objectionable about stereotypes in general (Blum 2004: 256).

If stereotypes were always false or misleading, one could diagnose what is epistemically wrong with implicit bias in simple terms. Implicit biases would be constituted by group stereotypes. Once internalized, these false claims would cause individuals to form inaccurate beliefs about social groups and the individuals that belong to them.

Thinking about the metaphor of bias as fog thus leads us to think of the epistemic significance of implicit bias in a particular way. Biases are always epistemically bad, if we adopt the metaphor.

Their badness is multi-dimensional. Biases are widely thought to articulate false or misleading claims about groups, which—once internalized—taint perceptual and cognitive judgments about individuals.

#### 4.1 Why Implicit Bias is Not Like A Fog

I don't think that implicit bias is a fog—at least not exactly. To see why, think about what Camp calls *centrality*. The metaphor of fog attributes key causal power to stuff outside of people's minds. The fog of oppression infects the social landscape, and it corrupts individuals' perception and cognition. This feature of the environment, if we take the metaphor seriously, is also beyond human control. We couldn't get rid of fog, even if we wanted to.

Focusing on the causal implications of the metaphor, we begin to see its problems. For example, in San Francisco, fog rolls in during the morning, sometimes leaves, and appears again at dusk. No one can stop this progression of things. In contrast, stereotypes are a *human* creation. Conditions of domination and oppression—and the ideologies that support them—are too. Smog is clearly the better metaphor here. Smog is “fog or haze intensified by smoke or other atmospheric pollutants” (OED 2017). Pollution is the result of human activity. It is something that humans could get rid of, in principle; however, doing so would be extremely difficult, given the tenacity of vested interests.

Modifying the metaphor in this way helps. Thinking of implicit bias as “the smog that hangs over a community”—rather than fog—captures why breathing it in would be bad for you

(Vedantam 2017). Smog causes bodily dysfunction, including lung damage, asthma, and other respiratory ailments. Cognitive smog causes perceptual and cognitive dysfunction. Thinking of implicit bias as smog also indicates a kind of a responsibility for the problem, both at the collective and individual levels. Human agency is implicated in implicit bias because people often actively stereotype and, also, actively ignore the ways in which they benefit from and contribute to unjust social conditions. Moreover, humans created the oppressive practices, laws, and ideologies that generate stereotypes in the first place.

Of course, the metaphor of bias as smog has its problems too. Smog primarily affects cities, and it is typically episodic. Biases are, however, everywhere, and they are unrelenting. There is no escaping bias by retreating to a rural existence or driving out of town. Perhaps the biggest challenge to the metaphor, however, emerges when we consider a competing metaphor: implicit bias as shortcut.

## 5. Taking Short Cuts

If one were looking for the most popular metaphor about bias, there would be no contest. Implicit bias is most often thought of as a shortcut (Ross 2014; Google 2014; UCLA 2016). In *The Times* video, one finds this metaphor alongside bias as fog. Yet the convergence is puzzling. The two metaphors have contradictory implications when it comes to the epistemic significance of bias. They also potentially diverge in their appraisals of *when* and *why* implicitly biased judgments undermine knowledge.

To see this, think about what a shortcut is. Here are two definitions: “a path or a course taken between two places which is shorter than the ordinary road” and “method of attaining some object” (OED 2017b).

The metaphor of bias as a shortcut is largely due to psychologists Daniel Kahneman and Amos Tversky. Since the early 1970s, their work on heuristics and biases has been enormously influential in psychology, economics, legal theory, and philosophy (Kahneman & Tversky 1973a; 1973b; 1974). Humans, they argue, often engage in *fast* ways of thinking. Fast thinking saves time and mental energy. It also sometimes results in correct predictions and can be reasonable. However, fast thinking is unreliable in certain contexts, leading “to severe and systematic errors” (Tversky & Kahneman 1973a: 237). Their life’s work consists in documenting the myriad of ways in which biases cause unreliable judgments.

In the 1980s and 90s, Kahneman & Tversky’s work was taken up by social psychologists who studied stereotyping. In an influential textbook on social cognition, Susan Fiske and Shelley Taylor wrote that humans are “cognitive misers” (Fiske & Taylor 1984). We have limited time, knowledge, and attention. Because of this, they argue, humans automatically opt for quick, efficient ways of thinking. Hence we stereotype. Stereotyping is a substitute for more careful, slow ways of forming judgments about individuals.

Embodied mind views of implicit bias also deploy the metaphor of bias as shortcut—and in a helpfully concrete way. Shortcuts are literally faster ways of getting from point A to point B. When I jump to conclusion about someone based on group stereotypes, I may be engaging in the

embodied activity of judging that person from afar, instead of the embodied activity of approaching them and getting to know them.

To see how biases function as shortcuts, consider an example that I have used elsewhere, which I call *I Need a Doctor* (Beeghly 2015). Imagine a panicked father in an emergency room. “Where is the doctor?” he might yell, “I need a doctor.” The man might grab the first person he sees in a white coat, relying on the stereotype that doctors wear white coats, not caring that he is grabbing this or that particular doctor, not caring about the doctor at all in his or her individuality. The father is using a heuristic to judge who is and who is not a doctor. The heuristic is this: white coat in an emergency room = doctor. This shortcut will sometimes work. However, it will sometimes also fail. For example, during a recent emergency room visit, I saw a sign on the wall. It read: “Doctor wears blue scrubs.” A sign like this was necessary because white coats are strongly associated with doctors. Stereotypically, doctors wear white coats. As one M.D. puts it, the white coat “has served as the preeminent symbol of physicians for over 100 years” (Hochberg 2007: 310). Likewise, white coats are associated with *competence*. As one recent study found: “patients perceived doctors as more trustworthy, responsible, authoritative, knowledgeable, and caring in white coats” (Tiang et al. 2017, 1). When the father expects that the person in the white coat is a competent doctor, he is thus doing something entirely typical. He is using a stereotypic association to identify someone as a doctor, and he is forming expectations about that person on that basis. Depending on how one describes the details of this case, his judgment may even count as a manifestation of *implicit* bias. However we describe the case, this much is clear: his judgment and behavior betrays reliance on a cognitive shortcut.

Remember, once again, that metaphors are supposed to do two important things, according to Camp. First, they render certain aspects of a phenomenon more *salient*, and they assert claims about the *centrality* of properties associated with the phenomenon.

What is rendered salient if we think of implicit biases as shortcuts? Their epistemic virtues! Shortcuts are, as per the definition offered earlier, “compendious,” which means “economical,” “profitable,” “direct,” and “not circuitous” (OED 2017c). To call stereotypes shortcuts is thus to pay them a compliment. It is to call attention to their pragmatic and cognitive usefulness—perhaps even their necessity.

This metaphor also emphasizes the universality of bias. Philosopher Keith Frankish writes:

an implicitly biased person is one who is disposed to judge others according to a stereotyped conception of their social group (ethnic, gender, class, and so on), rather than by their individual talents (Frankish 2016: 24).

Since all humans allegedly have the disposition to stereotype—because we all have cognitive, informational, and practical limitations and our world is complex, we are *all* implicitly biased. In *The Times* videos, Reshamwala emphasizes the normalcy and universality of bias repeatedly. “If you’re seeing this,” he says, “and are thinking that it doesn’t apply to you. Well, you might be falling prey to *the blindspot bias*. That’s a scientific name for a mental bias that allows you to see biases in others but not yourself. We’re [all] biased!”

One can already see how the metaphor of bias as shortcut differs from that of bias as fog. When someone says implicit bias is fog, they are committed to saying that it is *always* an obstruction, something that makes it harder to perceive and judge individuals clearly. When someone says that bias is a shortcut, they are saying that biases facilitate perception and judgment by providing an efficient means of judging and making predictions about individuals. These are not exactly diametrically opposed suggestions. But they emphasize different aspects of implicit bias. However, you do get a direct contradiction out of the metaphors if you pursue the idea of bias as a shortcut further.

Here is why. If implicit biases are shortcuts, they will sometimes yield “reasonable” and correct judgments about individuals. Every writer cited a few paragraphs back says as such, as do Kahneman & Tversky. Such claims are inconsistent with the claim that bias is fog. Indeed, they suggest that implicitly biased judgments sometimes let you cut right to the chase and can thus be epistemically good in some cases. In hospitals where doctors wear white coats, for example, and hospital staff wears other garb, relying on the stereotype of doctors as wearing white coats will help you quickly predict who is and who is not doctor.

### 5.1 The Diversity of Epistemic Objections to Bias

One helpful thing about the metaphor of bias as shortcut is that it invites us to think more carefully about the conditions under which implicitly biased judgments are epistemically problematic.

Consider, first, the objection that implicit biases are constituted by *false, unwarranted* stereotypes. As I noted in section 4, stereotypes are typically thought of as false or misleading group generalizations. Often they are also thought to be unwarranted by evidence. This way of thinking about stereotyping fits perfectly with the metaphor of bias as fog. However, once one starts to think of stereotypes as shortcuts, one begins to wonder, “is it really true that stereotypes are *always* false and based on limited experience?” The answer to this question is likely *no*.

Think about the following gender stereotype: women are empathetic. This stereotype is likely true, if considered as a claim about *most* women or as a claim about the relative frequency of empathic characteristics in women compared to men. We live in a patriarchal society. When women in a society like ours are raised to value empathy and actually tend to self-describe as empathic, when they tend to fill social roles where empathy is required or beneficial and when the social benefits of displaying empathy disproportionately go to women, women will, on average, have a greater disposition for empathy than men and one that is stable over time (Klein & Hodges 2001; Ickes 2003). Accordingly, the claim that women are empathetic could be true. Moreover, as feminist scholars have argued about similar stereotypes, we would be justified in implicitly or explicitly believing it was true (de Beauvoir 1953: xxiv; Haslanger 2012: 449; Haslanger 2017: 4).

Observations such as this complicate epistemic evaluations of bias. If biased judgments were always based on false, unwarranted beliefs about groups, we would have a decisive epistemic objection to people using them. Of course one shouldn't deploy false, unwarranted beliefs to judge individuals. On the other hand, if the stereotypes that drive biased judgments might

sometimes be true and warranted by the evidence, one cannot always invoke this objection to explain why people should *never* judge others in implicitly biased ways. After all, the objection will only *sometimes* apply. To find an objection that always applies, one must get more creative.

Thinking of biases as shortcuts helps here. In the literature on heuristics and biases—where the metaphor that we are considering originated—authors tend to articulate epistemic objections that apply to *processes* of reasoning that involve stereotyping. Consulting this literature, one finds ample reason to think that implicitly biased judgments are always or usually *unreliable*. The reason for their unreliability is not premised on the falsity or lack of justification of stereotypes. Biased judgments would be unreliable, according to these theorists, even if the stereotypes being deployed were true and warranted.

Here are three examples.

(A) *The Representativeness Heuristic*. Suppose someone handed you the following character sketch:

Steve is shy and withdrawn, invariably helpful but with little interest in people or in the world of reality. A meek and tidy soul, he has need for order and structure and a passion for detail (Kahneman 2011: 7).

That person then asks you, “Is it more probable that Steve is a librarian or a farmer?” What would you say?

If you were like typical research participants, you would say that Steve is probably a librarian. In giving this answer, one relies on what Tversky and Kahneman call *the representativeness heuristic*. Here is the *OED* definition of a heuristic, as understood by psychologists: “designating or relating to decision making that is performed through intuition or common sense” (*OED* 2017d). The opposite of heuristic is “systematic.” Systematic ways of reasoning adhere to the norms of ideal rationality, as modeled by decision theorists.

When people use the representativeness heuristic, they make judgments about the likelihood of people having this or that property—for example, the property of being a librarian—based on stereotypes. Thinking quickly, we automatically expect that Steve will be a librarian because he fits the stereotype of a librarian.

The problem with using the representativeness heuristic is that it involves ignoring a great deal of other information. “Did it occur to you,” writes Kahneman,

that there are 20 male farmers for each librarian in the United States? Because there are so many farmers, it is almost certain that more meek and tidy souls will be found at tractors than at library desks (7).

If you stereotyped Steve, he says, you committed *base rate neglect*. A person neglects base rates if they ignore background statistics—such the percentages of librarians and farmers in the population at large—when reasoning. Implicitly biased people, the criticism goes, always make

judgments by ignoring base rates. Their predictions and expectations of individuals will thus be unreliable.

(B) *The Availability Heuristic*. Implicitly biased people also make use of the *availability heuristic*. When people use this heuristic, Kahneman says, their task is to estimate the size of a category or the frequency of an event but ... [they instead] report an impression of ease with which instances come to mind” (130). Because one is not paying attention to actual probabilities, one ends up overestimating or underestimating the probability of an event or property occurring (Tversky & Kahneman 1973b; Lichtenstein et al 1978). This effect very often occurs when properties are dangerous or striking. But it may occur in other cases as well. The mere existence of a trait as part of a cultural stereotype may bring it more easily to mind than would otherwise be the case. For example, we may overestimate the percentage of mothers among women because, stereotypically, women bear children.

If implicitly biased people use the availability heuristic, they would often have unreliable predictions, expectations, educated guesses, and beliefs about individuals.

(C) *The Affect Heuristic*. Implicit biases may also leave us open to non-cognitive biases. Stereotypes can bring to mind aversions and affinities and are often laden with evaluative and emotional significance (Madva & Browstein 2016).

Some of the most interesting work on affect and biases has been done by Paul Slovic and colleagues. Slovic introduced the idea of an *affect heuristic*. As before, the idea with heuristics is

that people aim to find easy ways of answering questions when thinking quickly and intuitively. Emotions can be helpful for this purpose. A person may simply consult his feelings to determine what he should think and do. If one's feelings are clear cut, one can "just go with it" and suppose that affect provides the right answer to the question. "Using an overall, readily available affect impression can be easier and more efficient than weighing the pros and cons of various reasons or retrieving relevant examples from memory," writes Slovic, "especially when the required judgment or decision is complex or mental resources are limited" (Slovic et al. 2004: 314).

Think, first, about the content of stereotypes. Stereotypes will often be affectively laden. Gay men are stereotyped as pedophiles, for example. Muslims are stereotyped as violent extremists. Black men are stereotyped as criminals. Just as people vastly overestimate the likelihood of being attacked by a shark while swimming due to fear, they may vastly overestimate the likelihood that individuals from groups mentioned above will possess the dangerous properties stereotypically attributed to them. Fear—not facts—would guide estimation of probabilities. Of course emotions—especially ones like fear—are not a reliable sources of probabilistic information. So using this heuristic in conjunction with stereotypes would lead to unreliable judgments.

A second observation concerns relationship between moods, quick thinking, and stereotyping. What psychologists have found is that people in happy or positive moods often think quickly, hence they tend to stereotype (Park J. & M. Banaji 2000; Chartrand, Van Baaren, & Bargh 2006; Forgas 2011; Holland et al. 2012). For example, Forgas (2011) asked research participants to read a one-page philosophy essay written by "Robin Taylor." Attached to the essay was either a

picture of a middle-aged white man with glasses—a stereotypical-looking philosopher—or a young white woman with “frizzy” hair—someone who poorly fits the stereotype of a philosopher. When the essay was attributed to the middle-aged white male, participants tended to rate it more positively. This bias was most pronounced when participants were in a good mood. In contrast, participants in bad moods were less influenced by stereotypes when evaluating the essay. They spent more time reading and thinking about the essay, and they evaluated the essay as just as good no matter who wrote it. As such experiments show, affect plays a complex role in our epistemic life and can undermine our ability to evaluate others in fair, unbiased ways (Madva forthcoming).

By paying attention to the literature on biases and heuristics, we seem to have found a new epistemic objection to biases and one that allegedly applied in every single problematic case of bias. The objection is that biased judgments are unreliable because they are the product of fast thinking. What we need to do, the argument goes, is to slow down, reason more carefully, and judge persons as individuals.

Have we now found a fool-proof objection to bias? Perhaps not. Within the literature on heuristics and biases, theorists do not make the above argument. What they say is that fast thinking leads us astray *sometimes*. At other times, it is supposed to “yield reasonable judgments” (Kahneman & Tversky 1973a: 48; Jussim 2012: 360-388). If there are contexts in which stereotyping is reliable and reasonable—such as, perhaps, in *I Need a Doctor*—the challenge becomes how to distinguish the conditions under which stereotyping helps us track truth and those in which it leads us into severe and systematic error.

Here is another problem with the above argument: it appeals to the claim that we have an epistemic and moral duty to judge persons as individuals. However, it is not clear that such an epistemic duty exists or, if it does, how to articulate it. Philosopher Benjamin notes:

Taken literally, the principle [of treating persons as individuals] seems to express a broad hostility to forming judgments about individual people by appeal to generalizations about whole classes of people (Eidelson 2013: 204).

Understood in the way, it is absurd to think that there is an epistemic obligation to treat persons as individuals (Levin 1992: 23; Schauer 2006: 19; Arneson 2007: 787). Even if it were possible to refrain from judging others based on real or apparent group membership, the effects of doing so would be epistemically devastating. We would lack schemas for organizing our social world. We couldn't learn about groups of people. We would be forbidden from categorizing unfamiliar individuals as members of types and forming expectations about them based on group membership. For example, in *I Need a Doctor*, the father would be forbidden on epistemic grounds from identifying the white-coated person as a doctor.

The injunction to always treat persons as individuals in the way specified above is not only epistemically odd; it is also ethically troubling. Imagine a woman who believes that people of color in her community are often subject to police harassment. When she sees a random black man, she expects that he, too, has likely experienced police harassment at some point in his life. This person is “using race as a proxy for being subject to unjust race-based discrimination,” as

Elizabeth Anderson puts it (Anderson 2010: 161). This woman expects that a man—whom she has never met—is likely to have experienced police harassment simply because he is a black man. She is using a shortcut and is failing to treat the man as an individual. Yet, I would say, she has done nothing epistemically or ethically wrong. Indeed, this stereotyping may be the best possible ethical and epistemic response.

Perhaps the epistemic and moral duty to treat person as individuals can be interpreted in a more plausible way. For example, Kasper Lippert-Rasmussen has suggested:

X treats Y as an individual if, and only if, X's treatment of Y is informed by all relevant information, statistical or non-statistical, reasonably available to X (Lippert-Rasmussen 2011: 54).

Call this *the use-all-your-information conception* of treating persons as individuals. Adopting this conception, one might argue that implicitly biased people are always failing to treat persons as individuals because they fail to use all relevant, reasonably available information when judging others.

Such a claim fails, however. Implicitly biased judgments will only sometimes involve failing to treat person as individuals, as defined above. When biased agents face serious informational deficits—and thus have very little information reasonably available to them—they will count as treating persons as individuals, even though they stereotype others based on group generalizations (Beeghly ms). Likewise, it is possible for a biased agent to use all the relevant

information reasonably available to her in forming a prediction about someone; yet implicit group stereotypes—which she might disavow—may corrupt how she interprets and acts on that information. In such a case, there is something epistemically wrong with her judgment; however, it is not captured by the claim that she has failed to use all her information when reasoning. What’s gone wrong is something quite different. For example, it could be that she uses all her information but weighs different pieces of evidence inappropriately.

What is the upshot? Perhaps it is that the range of epistemic objections to implicit bias is astoundingly wide. Or, maybe the lesson here is that no single epistemic problem will be present in all epistemically problematic cases of implicitly bias.

The last possibility is very important to philosophers. One thing that philosophers like to do is create *theories*. A theory of what’s epistemically wrong with implicit bias could be unified or non-unified. *Unified theories* are so-called because they identify a single property or set of properties that all epistemically bad cases of bias allegedly have in common, in virtue of which the cases count as bad. *Non-unified theories* are so-called because they identify multiple properties that bad cases of bias might have in common. Though all wrongful cases are alleged to share one of many wrong-making properties specified in the list, no single property mentioned on the list will be found in every single wrongful case.

If the analysis so far is on the right track, we should not expect a unified theory of what’s epistemically wrong with implicit bias to succeed. Certainly implicit biases hamper our knowledge in many cases, but there seems to be no single objection that fully explains why they

do so in every case. Just as importantly, we have not yet been able to definitively rule out the possibility that implicitly biased judgments are sometimes epistemically rational and are, perhaps even, sometimes unobjectionable from an epistemic point of view. An important project for future research is to consider these issues more systemically, in the hopes of better understanding the conditions under which implicitly biased judgments are epistemically problematic.

## 5.2 Why Implicit Biases are Not Just Shortcuts

I don't believe that implicit biases are best understood as shortcuts. So I am not onboard with this second metaphor either, despite its advantages. There are a few reasons why.

First, thinking of bias as a shortcut encourages us to believe that biased judgments are primarily due to quick thinking. If that were true, we could rid ourselves of biases by attending more carefully to the facts. Yet, as philosopher Louise Antony notes,

...it is a kind of fantasy to think that biases intrude only when our guard is down—a fantasy that permits us to think that if we were only more careful in our thinking, more responsible or more virtuous in our epistemic practice, things would be all right. That leaves intact the conviction that there is within each one of us some epistemic still place from which we can see clearly and judge soundly... (Antony 2016: 160).

The reality is that humans are always, already biased, and there is no bias-free epistemic alternative. Biases affect what we pay attention to and ignore, what we see and don't see, what we remember and forget, and how we interpret evidence. Even when thinking carefully, biases

can shape our judgments. For example, even when employers evaluate job applications carefully, their biases can affect which candidates they prefer and how they evaluate credentials (Rivera & Tilcsik 2016; Dovidio & Gaertner 2000). A recent meta-analysis of the role of gender in hiring decisions, for example, found that people who were motivated not to discriminate displayed *less* gender bias when evaluating women job candidates in male-dominated fields; however, they were still not able to get rid of their biases completely (Koch et al. 2015).

Second, by using the metaphor of a shortcut, one implies that biased people are holding and using stereotypes because they are lazy, rushed for time, or overwhelmed by the world's complexity (Bargh 1999). However, as Walter Lippman noted in 1922, stereotypes “are not merely short cuts” but are a “projection upon the world of our own sense of our own value, our own position and our own rights” (Lippman 1922/1997: 63-64). In other words, implicitly biased judgments don't just occur because people are pressed for time and overwhelmed with stimuli; they happen because we exist in a world where certain kinds of people stand in particular power relationships to one another. Stereotyping—whether implicit or explicit—is always wrapped up in power, privilege, ideologies, and histories of oppression.

It is no accident that the metaphor of bias as shortcut largely hides the ideological and social dimensions of bias. The metaphor identifies implicit biases with mental states; hence it renders the psychological elements of the phenomenon central and salient. Biases are typically described as *cognitive* shortcuts, after all. Nothing is said about their origin, and their existence is often alleged to be a matter of innate cognitive architecture. The connection between bias, the social

world, and group oppression—which was foregrounded in the metaphor of bias as fog—is thus lost.

## 6. Concluding Thoughts on the Epistemic Significance of Implicit Bias

In this chapter, I have explored two metaphors used to think about implicit bias: bias as fog and bias as shortcut. Like any metaphor, neither one is perfect. Both misrepresent the phenomenon of bias in some respects. On the other hand, both metaphors bring something important to the table. Perhaps that is why the two co-exist in the Reshamwala's video for *The New York Times*. Thinking of bias as a shortcut encourages us to pay attention to the relationship between biased judgments and fast thinking. Implicit biases do, in fact, often manifest when we think quickly. Thinking of bias as fog, in contrast, brings out its connection to group oppression, and its epistemic costs.

Considering these two metaphors together is also productive for a different reason: it calls attention to a pressing question about the epistemic significance of bias. That is, what is the actual connection between bias, knowledge, and error? If implicit bias is fog, it is always epistemically bad; however, if biases are shortcuts, implicitly biased judgments are not always bad from the point of view of knowledge. Which claim is correct?

We are still not in a position to answer this question, which is significant. The crux of the matter is whether implicit bias has any epistemically positive role to play in our individual and collective attempts to gain knowledge of the world and of the people in it. If one could show that

bias is epistemically productive and that there are even cases in which there is nothing epistemically wrong with biased judgments, there would be little reason to think of it as fog or smog, i.e., something that necessarily harms our ability to clearly see and make judgments about the world.

The argument in favor of thinking of bias as epistemically productive goes as follows. Implicitly biased judgments are the result of stereotypes. We could not be fully rational but for stereotypes and stereotyping. Stereotypes help us cope with deeply human limitations. When we try to learn things about the world and other people, we have scarce resources with which to work: limited time, attention, intellectual capacities, and perspective. Relying on group generalizations helps us to overcome these limitations. It enables predictions about unfamiliar individuals with whom we've had little interaction. It also obviates the need to see every individual as completely unique, thereby freeing up our attention and cognitive energy for the more pressing tasks. If we were different kinds of creatures, we could, perhaps, judge every person as an individual and wouldn't have to rely on kind-based thoughts and reasoning; but, as things are, we need stereotyping to lighten our cognitive load and to guide us in conditions of limited information. Moreover, even though we use them quickly, stereotypes can be based on a lifetime of experience, perception, even wisdom. Given all this, the argument goes, it is highly implausible that any and all instances of stereotyping—and, by extension, all implicitly biased judgments—are worthy of epistemic condemnation.

The argument picks up further steam when one considers that stereotyping—the alleged cause of implicitly biased judgments—shares a good deal in common with inductive reasoning about

kinds of things in general. Notice that we have “pictures in our heads” of lightening storms and rivers, tables and skyscrapers, skunks and otters, just as we have stereotypes of social groups. We make generalizations about all kinds of things, and doing so is epistemically useful. By relying on kind-based generalizations, we save time and energy. We get around better in the world, having a better sense of what to expect from new things, situations, and people. We can avoid potentially dangerous situations and seek out advantageous ones. We can make reliable predictions. Stereotyping can also fail in all the same ways as kind-based reasoning more generally. We may form stereotypes based on a limited sample size then overgeneralize. Our past experience with social groups may not be a reliable guide to the future. Our expectations can lead us to pay attention only to what confirms them and ignore disconfirming evidence. We may systemically over-estimate the likelihood of events based on heuristics. Despite these problems, no one is tempted to say that kind-based reasoning in general is always epistemically bad.

The above claims culminate in what I call *the argument from symmetry*. The argument goes like this. If we claim that it is always epistemically bad to use stereotypes (which is what happens when people make implicitly biased judgments), we will have to endorse this thought in other domains as well. For example, we will have to say that there is always something epistemically bad with forming expectations about objects like chairs or nonhuman animals or physical events like lightning storms on the basis of group membership. Yet, the argument continues, it is very hard to believe that kind-based reasoning about *anything whatsoever* is necessarily epistemically problematic. So, the argument goes, why would kind-based reasoning about people always be epistemically problematic? Parity of reasoning requires us to see stereotyping people as

sometimes rational and, indeed, as potentially good from the perspective of knowledge *sometimes*.

If one is to defend the epistemic claim behind the metaphor of implicit bias as fog, namely, that there is always something epistemically wrong with biased judgments, one must push back against the argument from symmetry. One strategy is this: identify special epistemic problems that occur when we deploy social stereotypes in perception and cognition, which do not occur when we use other kinds of generalizations. A second strategy is to argue that we face higher epistemic standards when judging persons (and, perhaps, non-human animals) for ethical reasons and, thus, lower epistemic standards apply when we reason about other kinds of things (Beeghly 2014: 79-80; Basu cf. this volume). A third strategy is to endorse symmetry but argue that there is always something epistemically problematic with kind-based reasoning, no matter what we are reasoning about. Whether any of these strategies works is a source of continued debate.

I am not sure which metaphor will ultimately win out. Both are tempting, albeit for different reasons. But, more than anything, what they reveal is how much there is to learn about the conditions under which biased judgments are epistemically problematic. This essay has only been able to scratch the surface. Investigating further, one would have to consider a range of other epistemic objections to biased judgments, actions, and speech (Blum 2004; Gendler 2001; Haslanger 2012; Medina 2013). Even just scratching the surface, one can understand why these two metaphors are so prevalent in popular discussions of implicit bias: they simplify and provide an accessible frame from which to begin deeper philosophical reflections on the nature of implicit bias and its connection to social injustice. They also bring us to the heart of a deep

disagreement about the epistemic significance of implicit bias, a disagreement that currently has no end in sight.

## REFERENCES

- Anderson, E. (2010) *The Imperative of Integration*, Princeton, NJ: Princeton University Press.
- Antony, L. (2016) "Bias: Friend or Foe? Reflections on Saulish Skepticism," in M. Brownstein & J. Saul (eds.) *Implicit Bias and Philosophy: Volume 1*, Oxford: Oxford University Press, 157-190.
- Arneson, R. (2007) "What is Wrongful Discrimination?" *San Diego Law Review*, 43: 775-808.
- Banaji M. & Greenwald, A. (2013) *Blindspot: The Hidden Biases of Good People*, New York, NY: Delacorte.
- Bargh, J. (1999) "The Cognitive Monster: The Case Against the Control of Automatic Stereotype Effects," in S. Chaiken & Y. Trope (eds.) *Dual-process Theories in Social Psychology*, New York, NY: Guilford Press, 361-382.
- Beeghly, E. (2014) *Seeing Difference: The Ethics and Epistemology of Stereotyping*, PhD Dissertation, University of California, Berkeley, California.
- Beeghly, E. (2015) "What is a Stereotype? What is Stereotyping?" *Hypatia*, 30: 675-691.
- Beeghly, E. *What's Wrong with Stereotyping?* Manuscript in progress.
- Blum, L. (2004) "Stereotypes and Stereotyping: A Moral Analysis," *Philosophical Papers*, 33: 251-289.
- Camp, E. (Forthcoming) "Imaginative Frames for Scientific Inquiry: Metaphors, Telling Facts, and Just-So Stories," in P. Godfrey-Smith & A. Levy (eds.) *The Scientific Imagination*, Oxford: Oxford University Press.
- Chartrand, T, R. van Baaren, & J. Bargh (2006) "Linking Automatic Evaluation to Mood and Processing Style: Consequences for Experienced Affect, Impression Formation, and Stereotyping," *Journal of Experimental Psychology*, 135: 70-77.
- Dovidio, J. & S. Gaertner. (2000) "Aversive Racism and Selection Decisions: 1989 and 1999," *Psychological Science*, 11: 315-319.

- Dasgupta, N. (2013) "Implicit Attitudes and Beliefs Adapt to Situations: A Decade of Research on the Malleability of Implicit Prejudice, Stereotypes, and the Self-Concept," *Advances in Experimental Social Psychology*, 47: 233-279.
- Eidelson B. (2013) "Treating People as Individuals," in D. Hellman & S. Moreau, *Philosophical Foundations of Discrimination Law*, Oxford: Oxford University Press, 203-227.
- Fiske, S. & S. Taylor (1984/1991) *Social Cognition*, 1st ed, New York, NY: McGraw Hill Inc.
- Forgas, J. (2011) "She Just Doesn't Look Like a Philosopher...? Affective Influences on the Halo Effect in Impression Formation," *European Journal of Social Philosophy*, 41: 812-817.
- Frankish, K. (2016) "Playing Double: Implicit Bias and Self-Control," in M. Brownstein & J. Saul (eds.) *Implicit Bias and Philosophy: Volume 1*, Oxford: Oxford University Press, 23-46.
- Fricke, M. (2009) *Epistemic Injustice: Power and The Ethics of Knowing*, Oxford: Oxford University Press.
- Funchess, M. (2014) "Implicit Bias—How It Affects Us & How We Push Through" *Tedx Talk*, [ONLINE], 16 October 2014. Available at <https://www.youtube.com/watch?v=Fr8G7MtRNlk> [Accessed: 10 June 2017].
- Gendler, T. S. (2011) "On The Epistemic Costs of Implicit Bias," *Phil Studies* 156: 33-63.
- Google. (2014) "Google Video on Implicit Bias: Making the Unconscious Conscious," [ONLINE], 25 September 2014. Available at [https://www.youtube.com/watch?v=NW5s\\_-NI3JE](https://www.youtube.com/watch?v=NW5s_-NI3JE) [Accessed: 10 June 2017].
- Haslanger, S. (2012) "Ideology, Generics, and Common Ground," *Resisting Reality: Social Construction and Social Critique*, Oxford: OUP.
- Holland, R., M. de Vries, B. Hermsen, & A. van Kippenberg. (2012) "Mood and the Attitude-Behavior Link: The Happy Act on Impulse, The Sad Think Twice," *Social Psychology and Personality Science* 3: 356-264.
- Hochberg, M. (2007) "History of Medicine: The Doctor's White Coat—A Historical Perspective," *American Medical Association Journal of Ethics*, 9: 310-314.
- Ickes, W. (2003) *Everyday Mind Reading: Understanding What Other People Think and Feel*, Amherst, NY: Prometheus.
- Jussim, L. (2012) *Social Perception and Social Reality*, Oxford: Oxford University Press.
- Kahneman, D. & A. Tversky (1973) "On the Psychology of Prediction," *Psychological Review*, 80: 237-251.

\_\_\_\_\_. (1973b) "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 5: 207–232.

\_\_\_\_\_. (1974) "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185: 1124-1131.

Kahneman, D. *Thinking Fast and Slow*. New York, NY: Farrar, Strauss, & Girroux, 2011.

Kang, J. (2016) "Implicit Bias, Preface: Biases and Heuristics," *Bruin X*, UCLA Office of Equity, Diversity, and Inclusion [ONLINE], September 9, 2016. Available at: <https://www.youtube.com/watch?v=R0hWjJfDyCo&feature=youtu.be> [Accessed: 10 June 2017].

Klein, K. & S. Hodges (2001) "Gender Difference, Motivation, and Empathic Accuracy: When It Pays To Understand," *Personality and Social Psychology Bulletin* 27: 720-730.

Koch A., S. D'Mello, & P. Sackett. (2015) "A Meta-Analysis of Gender Stereotypes and Bias in Experimental Simulations of Employment Decision Making," *Journal of Applied Psychology*, 100: 128-161.

Lakoff J. & Mark Johnson. (1980) *Metaphors We Live By*. Chicago, IL: University of Chicago.

Levin, M. (1992) "Responses to Race Differences in Crime," *Journal of Social Philosophy*, 23: 5-29.

Lichtenstein et al. "Judged Frequency of Lethal Events," *Journal of Experimental Psychology: Human Learning and Memory* 4 (1978): 551-578.

Lippert-Rasmussen, K. (2011) "We are all Different": Statistical Discrimination and the Right to be Treated as an Individual," *Journal of Ethics*, 15: 47-59.

Lippman, W. (1922/1997) *Public Opinion*, New York, NY: Free Press.

Madva, A. (Forthcoming) "Implicit Bias, Moods, and Moral Responsibility," *Pacific Philosophical Quarterly*.

Madva A. & M. Brownstein, "Stereotypes, Prejudice, and the Taxonomy of the Social Mind," *Nous*, 50: doi:10.1111/nous.12182.

Medina, J. (2013) *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*, Oxford: OUP.

Mills, C. (1999) *The Racial Contract*, Ithaca, NY: Cornell University Press.

*Oxford English Dictionary Online*. (2017a) "fog, n.2," Oxford University Press [ONLINE] March 2017. [Accessed: 31 May 2017].

———. (2017b) “shortcut \ short-cut, n.1,” Oxford University Press [ONLINE] March 2017. [Accessed: 31 May 2017].

———. (2017c) “compendious, adj,” Oxford University Press [ONLINE] March 2017. [Accessed: 31 May 2017].

———. (2017d) “heuristic, adj. B1b,” Oxford University Press [ONLINE] March 2017. [Accessed: 12 June 2017].

Park, J. & M. Banaji (2000) “Mood and Heuristics: The Influence of Happy and Sad States on Sensitivity and Bias in Stereotyping” *Journal of Personality and Social Psychology*, 78: 1005-1023.

Rivera, L. & A. Tilcsik. (2016) “Class Advantage, Commitment Penalty: The Gendered Effect of Social Class-Signals in an Elite Labor Market,” *American Sociological Review* 81: 1097-1131.

Ross, H. (2014) *Everyday Bias: Identifying and Navigating Unconscious Judgments in Everyday Life*, New York, NY: Rowman & Littlefield Publishers.

Reshamwala, S. (2016) “Peanut Butter, Jelly, and Racism” in *Who Me Biased? New York Times Video* [ONLINE], Available at: <https://www.nytimes.com/video/us/100000004818663/peanut-butter-jelly-and-racism.html> [Accessed: 10 June 2017].

Siegel, S. (2017) *The Rationality of Perception*, Oxford: Oxford University Press.

Slovic, P et al. (2004) “Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality,” *Risk Analysis*, 24: 311-322.

Tiang, K.W., A.H. Razack, & K.L. Ng (2017) “The ‘Auxiliary’ White Coat Effect in Hospitals: Perceptions of Patients and Doctors,” *Singapore Medical Journal*, [ONLINE] 4 April 2017. Available at: doi: 10.11622/smedj.2017023

Vedantam, S. (2017) “In the Air We Breathe,” *Hidden Brain Podcast*, National Public Radio, Episode 74, [ONLINE], 5 June 2017. Available at: <https://player.fm/series/series-1324366/in-the-air-we-breathe> [Accessed: 10 June 2017].

Weldon, M. (2016) “Implicit Gender Bias: Strategies to Own the Power To Succeed as Women Leaders,” *The Movement Blog* [ONLINE], 1 November 2016. Available at: <https://www.taketheleadwomen.com/blog/implicit-gender-bias-strategies-to-own-the-power-to-succeed-as-women-leaders/> [Accessed: 10 June 2017].

## ANNOTATED READING SUGGESTIONS

If you’d like to read more about how metaphors structure thought, read:

- Lakoff J. & Mark Johnson. (1980) *Metaphors We Live By*. Chicago, IL: University of Chicago. Lakoff and Johnson argue that metaphors are fundamental to how we understand and reason about the world and others. A classic text that has inspired a great deal of current research.

If you want to learn more about the metaphor of bias as shortcut, read:

- Kahneman, D. (2011) *Thinking Fast and Slow*. New York, NY: Farrar, Strauss, & Girroux,. Kahneman presents his and Amos Tversky’s research from the last four decades for a popular audience. He argues that stereotyping is not always bad from an epistemic perspective, while documenting the myriad ways in which people who rely on stereotypes as heuristics can be criticized on epistemic grounds.

If you are looking for an argument that bias—like fog—exists in environments and not just in the heads of individuals, read:

- Payne, K., H. Vuletich, & K. Lundberg. (in press) “The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice,” *Psychological Inquiry*.

If you’d like to read more about ways in which implicitly biased judgments might promote knowledge, see:

- Antony, L. (2016) “Bias: Friend or Foe? Reflections on Saulish Skepticism,” in M. Brownstein & J. Saul (eds.) *Implicit Bias and Philosophy: Volume 1*, Oxford: Oxford University Press, 157-190. According to Antony, “bias is an essential element in epistemic success” and “biases make possible perception, language, and science.”
- Beeghly, E. (2015) “What is a Stereotype? What is Stereotyping?” *Hypatia*, 30: 675-691. Beeghly distinguishes two ways to understand the terms “stereotype” and “stereotyping,” and she argues that stereotyping—understood as judging persons by group membership—is likely to be epistemically and ethically unproblematic in some cases.
- Jussim, L. (2012) *Social Perception and Social Reality*, Oxford: Oxford University Press. Jussim argues that stereotypes can be accurate and evidentially warranted; he also argues that forming judgments about individuals based on stereotyping is sometimes epistemically rational.

If you are looking to explore epistemic objections to biased judgments further, check out:

- Blum, L. (2004) “Stereotypes and Stereotyping: A Moral Analysis,” *Philosophical Papers*, 33: 251-289. Blum argues that stereotypes are necessarily false and resistant to evidence, and he claims that stereotyping will always be epistemically (and ethically) defective on three grounds, one of which is that stereotyping always involves failing to treat persons as individuals.
- Dotson, K. (2012) “A Cautionary Tale: On Limiting Epistemic Oppression,” *Frontiers*, 33: 24-47. Dotson argues that biased judgments—implicitly and explicitly biased

judgments—can cause epistemic oppression and epistemic exclusion, as well as undermine epistemic agency. A criticism of Fricker, see below.

- Fricker, M. (2009) *Epistemic Injustice: Power and The Ethics of Knowing*, Oxford: Oxford University Press. Fricker argues that implicitly biased judgments hinder knowledge by giving people less credibility than they are due—thereby constituting *testimonial injustice*—and by limiting people’s interpretative horizons—thereby constituting *hermeneutical injustice*. She also explores epistemic virtues that one might cultivate to avoid bias-generated injustices.
- Gendler, T. S. (2011) “On The Epistemic Costs of Implicit Bias,” *Phil Studies* 156: 33-63. Gendler discusses three epistemic costs of implicitly biased judgments, and she argues that some implicitly biased judgments might be epistemically rational but morally problematic.
- Haslanger, S. (2012) “Ideology, Generics, and Common Ground,” *Resisting Reality: Social Construction and Social Critique*, Oxford: OUP. Haslanger argues that pernicious stereotypes can be true and evidentially warranted, but using them in conversational contexts, e.g., saying things like “women are empathetic” is epistemically problematic because the utterances have false, unwarranted essentialist implications.
- Medina, J. (2013) *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*, Oxford: OUP. Medina explores the epistemic vices associated with biases, including close-mindedness, lack of imagination, and hubris. He also documents epistemic virtues the members of oppressed groups are likely to possess.

## DISCUSSION QUESTIONS

(We will pick 3 or 4 questions for the book & have the rest available online for instructors to use at their leisure.) \* = preferred discussion questions for book (EB)

What is an epistemic criticism? Give at least two examples from your own life; there is no need for the examples to be related to implicit bias.

\*What are the main pros and cons of the two central metaphors for bias considered in this chapter? Which one of the metaphors is most apt, in your view? Can you think of other, better metaphors for talking about bias?

Other metaphors for bias include "mirror-like reflections," "symptom of an underlying disease," "mind bugs," "collective hallucinations," "mental machinery", "mental setting or background," etc. Pick three of the above metaphors. What messages would each metaphor send about the nature or significance of implicit bias?

\*This chapter focuses on the question what makes implicit biases bad from the perspective of knowledge. Beeghly argues that there is likely no unified epistemic problem with biases. Why does she say this? For each potential epistemic problem with implicitly biased judgments, think of an example that shows that it doesn't apply in every case. Can you think of possible ways of unifying the epistemic objections, so as to create a unified theory of when and why implicitly biased judgments are epistemically problematic?

\*Even if a certain claim is partly warranted from an epistemic perspective (e.g., if most women are more empathic than most men), it might still, at least sometimes, be *morally problematic* to use that information in making judgments about other people (cf. Basu chapter). What are some potential ethical problems that might face the use of such biases—even if we are just focusing on the "biases" that seem to be relatively accurate?

Beeghly suggests that the fog metaphor threatens to undermine our ability to take responsibility for our biases (cf. responsibility chapter, structural chapter, intervention chapter). Do you think she is right? Why or why not?

\*Beeghly suggests that the use of stereotypes in the *I Need a Doctor* case is not obviously problematic from the perspective of knowledge. Do you agree? Can you think of other contexts where attitudes that look like problematic stereotypes might help support knowledge? When might they be valuable or necessary shortcuts, if at all?

Think of contexts where the Representativeness, Availability, and Affect Heuristics might *promote* knowledge and *hinder* knowledge. Think about whether there will be shared features that we can generalize from to think more broadly about when to put these heuristics into place as useful shortcuts, and when to put up structural or mental "fences" to block the usage of these shortcuts (like the blue-scrubs sign).

\*Ethically and epistemically, there is something very intuitive about the claim that we ought to treat people as individuals. However, many philosophers have argued that it is absurd to think that we must always treat persons as individuals, if that means never using group-based generalizations to form judgments and expectations about others. In your opinion, how should we interpret the ethical and epistemic demand to treat persons as individuals? Do you think it is possible that failing to treating persons as individuals could advance knowledge in some cases? Are some forms of generalization different from others, such that we think some generalizations fail to treat people as individuals (e.g., based on race or gender) and others don't (e.g., based on how people choose to dress or display themselves)?

### **LIST OF (POSSIBLE) GLOSSARY TERMS**

Epistemic objection

Unwarranted

Unreliable

Heuristic

Unified theory

Disjunctive/Non-unified theory

Generalization

Saliency

Centrality

### **WEB RESOURCES**

ONLINE VIDEOS

- Saleem Reshamwala, “Peanut Butter, Jelly, and Racism” in *Who Me Biased? New York Times Video*, December 2016. Available at: <https://www.nytimes.com/video/us/100000004818663/peanut-butter-jelly-and-racism.html>
- Google. “Google Video on Implicit Bias: Making the Unconscious Conscious,” September 6, 2014. Available at: [https://www.youtube.com/watch?v=NW5s\\_-NI3JE](https://www.youtube.com/watch?v=NW5s_-NI3JE)
- Jerry Kang, “Implicit Bias, Preface: Biases and Heuristics,” *Bruin X*, UCLA Office of Equity, Diversity, and Inclusion. September 9, 2016. Available at: <https://www.youtube.com/watch?v=R0hWjffDyCo&feature=youtu.be>

#### PODCASTS

- Shankar Vedentam, “In the Air We Breathe,” *Hidden Brain Podcast*, National Public Radio, Episode 74, June 5, 2017. Available at: <https://player.fm/series/series-1324366/in-the-air-we-breathe>

#### MOVIES

- J. Bush, R. Moore, & B. Howard, (2016) *Zootopia*, Hollywood, CA: Disney Studios.