Debiasing the Philosophy Classroom

LISA KRETZ
University of Evansville

Abstract: This paper is situated at the intersection of ethics, pedagogy, and bias. Various challenges for pedagogy that are posed by explicit and implicit bias are discussed. Potential solutions to such challenges are then explored. These include practices such as enhanced thought experiments, interviews, research projects, in-depth role-playing, action projects, and appropriately morally deferential experiential service-learning. Moral imagination can be beneficially stretched through adopting differing moral lenses and engaging and encouraging multiple empathizing; art and literary narrative provide helpful tools to this end. Also recommended is critical scrutiny focused on personal biases (including teacher bias) and developing curriculum focused on moral literacy. Such moves of necessity span from individual to public action given the environmental components of the operations of bias. Shaping ourselves through intentional environment construction and avoidance of undesirable environments is therefore identified as a valuable technique. Finally, the potential contribution of loving-kindness meditation is addressed. Although we may be unable at present to eradicate problematic forms of bias, there are multiple methods available to begin to ameliorate the harms associated with those forms of bias.

Introduction

Bias is starting to gain some much deserved notoriety in philosophical ethics. Although a rich body of literature about bias exists in psychology, thereby reflecting the importance of thinking about bias, bias has not yet taken center stage as a core issue in philosophical ethics, let alone in the teaching of ethics at the university level—which is the focus of this paper. This lack of a sustained study of bias as it pertains to academic ethics generally, including moral pedagogy, is surprising insofar as ameliorating bias often plays a crucial role in treating others with respect, equality, and compassion. This lack of sustained focus, however, is not so surprising when one realizes the quagmire of issues

that become apparent the moment one starts to consider the intersections of bias, ethics, and pedagogy. Such issues include the fact that the very definition of bias remains contested.

An interpretation that is common in the psychological literature takes bias to be "neutral with respect to both moral properties and questions of accuracy" (Kenyon and Beaulac 2014: 344). Understood as such, bias is "a disposition, implicit or explicit, to reach a particular kind of conclusion or outcome, or to remain in one" (Kenyon and Beaulac 2014: 344). Such a view accommodates the possibility that "biases can skew a process in a way that makes its outcome inaccurate or otherwise wrong, but it also leaves open the prospect that biases play a role in truth-conducive reasoning processes and morally unproblematic judgments or attitudes" (Kenyon and Beaulac 2014: 344). As such bias can be taken to a) have a positive or negative moral valence or remain neutral, and to b) have a positive or negative epistemic function or remain neutral.

In contrast, the more widely used colloquial understanding of bias does tend to have a negative valence both morally and epistemically. Although Tim Kenyon and Guillaume Beaulac accept the above definition, which leaves open the possibility of a negative, positive, or neutral valence to bias, they also acquiesce that one "normally goes to the trouble of saying that some attitude, reasoning, or person is biased only if the operation of the bias is claimed to be problematic—a distortion, or a prejudice that amounts to a vice" (Kenyon and Beaulac 2014: 344). If biases are evidence of prejudice then they are morally questionable, and if biases are distortions of what is actually the case they can negatively impact knowledge claims. I contend that a resolved definition of bias is not necessary for what follows to be of use to philosophy teachers. Whether one takes bias to be necessarily negatively valenced, or on occasion to be negatively valenced, insofar as it is the case that in a number of instances bias is morally and epistemically problematic, then bias should be of concern to professors of ethics. Although there is room for debate and further nuance regarding which biases are morally and epistemically problematic, in what follows I take there to be clearly problematic cases of bias. One example that will be discussed is implicit and explicit positive biases toward white people which serve to entrench systematic oppression.

The first section of this paper focuses on identifying various forms of bias, including some of the unique challenges these forms of bias present for teaching ethics. Both explicit and implicit forms of bias are highlighted. The second half focuses on developing solutions to such challenges through teaching practice. Various potential correctives to problematic operations of bias in the ethics classroom are explored. Although no method is adequate in isolation, a suite of potential ap-

proaches is identified so as to aid in working against biases. Even though working against the problematic operations of bias at multiple levels is no guarantee that a panacea has been found for the harm such biases can inflict on ethical knowledge and practice, it is only through these efforts that the operations of problematic biases can begin to be lessened insofar as we are capable of lessening them.

In closing, a novel method for working against both explicit and implicit bias is identified, namely loving-kindness meditation practice. Loving-kindness meditation serves to decrease bias through increasing controlled processing and decreasing automatic processing. Explicit and implicit bias can be simultaneously decreased through a seven-minute exercise that can be utilized in classes. Although we may be unable at present to eradicate bias, there are multiple methods available, operating at both explicit and implicit levels, to ameliorate problems associated with bias. We would do well, I argue, to make this a focus of academic study more widely and to apply the insights discovered to methods for teaching ethics.

Types of Bias[1]

One suite of biases concerns cognitive biases. Cognitive biases refer to ways in which thinking becomes biased, and there are a number of them. Specifically, with regard to moral issues, Keith Horton considers how four biases, illustrated through psychological testing, impact our ability to think about moral issues. One is belief perseverance, which refers to the tendency to maintain existing beliefs over time in part due to the tendency to seek out information that confirms one's existing beliefs and potentially neglecting contrary evidence to one's existing beliefs (Horton 2004: 548). Another is epistemic conformity, which refers to the tendency of our own beliefs to be strongly influenced by what we think others believe (Horton 2004: 549). This form of bias suggests significant moral change requires shifts in the social accept-ability of certain views and values. A popular example of epistemic conformity is bystander apathy, which illustrates a failure of people to help in emergency situations when others are not helping (Horton 2004: 548). An additional bias speaks to how vivid, concrete data is more influential than the same, or more troubling, data when it is pre-sented in a more abstract or pallid way (Horton 2004: 549). Here, part of the worry is that when immoral harms are suffered by others, but presented in an abstract or less vivid way, they will not be addressed. Another bias that can be at play is wishful thinking, which speaks to the tendency to be inclined to believe what we wish to be true (Horton 2004: 549). Wishful thinking points to one's bias in favour of a desired conclusion, so the information that supports a desired conclusion is as-

sessed less critically than information that does not support the desired conclusion (Horton 2004: 550).

Additional potential challenges lie in what Martin Hoffman calls empathetic biases. These include: familiarity bias, in-group bias, friendship bias, similarity bias, and here and now bias (Hoffman 2000: 207–12). While research shows that "most people empathize with and help others in distress, including strangers" there is also evidence that shows "most people empathize to a greater degree (their threshold for empathic distress is lower) with victims who are family members, members of their primary group, close friends, and people whose personal needs and concerns are similar to their own" (Hoffman 2000: 197). With regard to familiarity bias, when one is repeatedly exposed to any stimulus—even an inanimate one—a preference for that stimulus is developed (Hoffman 2000: 206). One form of familiarity bias is in-group bias.

> Evolution theorists agree that humans evolved in small groups and that although altruism was necessary for survival within groups, the scarcity of resources often pitted one group against another. It should therefore not be surprising that a person is more likely to empathize with and help those who are members of his or her family, ethnic or racial group—his or her in-group, in short. And when we consider that members of one's in-group are similar to each other and to oneself and share feelings of closeness and affection it should not be surprising that a person is also more likely to empathize with friends than with strangers and with people who are similar to oneself than with people who are different. (Hoffman 2000: 206)

Hence we have in-group bias, friendship bias, and similarity bias. Here and now bias involves tending to harmed persons who are present and immediately in the harmful situation (Hoffman 2000: 209). Empathy is "biased in favor of victims in the immediate situation because the automatic, involuntary, and salient empathy-arousing processes (conditioning, association, mimicry) require situational and personal cues that are at their peak when a victim is present" (Hoffman 2000: 209). Empathy-arousing processes are triggered by the immediacy of the situation. Hoffman notes "a cry of pain may arouse more empathic distress than a facial grimace; a friend's or relative's cry, more than a stranger's; the distress of someone who is present, more than that of someone who is absent" (Hoffman 2000: 197–98). Reflective awareness of how the above orientations function is thus necessary for an informed moral pedagogy.

Insofar as we as teachers and ethicists encourage our students to adopt perspectives that differ from their own, it will also be important to attend to how rhetorical credibility is established and the ways in which marginalized groups' rhetorical credibility is unjustifiably reduced (Yap 2013: 104). As Trudy Govier explains, rhetorical credibility refers to

the extent to which one is regarded as believable, and is believed, by others. People who are white and male, who dress well, look professional, appear middle or upper middle class, speak without an accent in a deep or low-toned voice, and seem unemotional, rational and articulate, tend in many contexts to have more rhetorical credibility than others. Often those who lack such qualities are, in effect, rhetorically disadvantaged. (Govier 1993: 94; also in Yap 2013: 104)

Some groups are implicitly rhetorically privileged (Yap 2013: 104). We are often unaware of our biases related to prejudices; for example, in our evaluations and expectations of others through gender schemas, as well as differential treatment due to race (Yap 2013: 105). If critical thinking is to help uproot unjustified biases, then it is pedagogically imperative to identify their operations (Yap 2013: 108). Multiple studies show that race and gender can skew evaluation of individuals, even though the evaluators would articulate that such considerations would not skew their position (Yap 2013: 105). Implicit bias, therefore, must be highlighted to our students so as to make explicit how it operates, and methods for ameliorating its problematic operations must be identified.

Implicit bias is a form of tacit bias. Bias can, of course, operate at more explicit levels as well. For example, someone can point out the implicit operations of my bias by pointing out my exclusionary actions toward certain groups (e.g., I tend not to socialize with Republicans), or I can actively recognize I have an explicit bias against certain groups (e.g., I have an explicit and intentional tendency not to socialize with Republicans). Such biases unfairly lump together broad collections of individuals, treating them as all having the same characteristics even though they do not.[2] These particular biases involve adherence to a problematically partial, or outright false, perspective on the nature of others and how we ought to respond to them. When bias functions implicitly, it poses uniquely knotty conundrums for the teaching and learning of ethics.

Deborah Mower (2015) invites scholars to engage with the challenge implicit bias poses to teaching ethics across the curriculum.[3] Mower helpfully illustrates a number of issues at the intersection of ethics, teaching, and practice in relation to implicit bias. She argues that whatever one's goals are with regard to ethics across the curriculum, part of what is assumed is that change is effected through our courses; unfortunately, however, there is reason to be less sanguine about current methods for teaching ethics and attendant pedagogical success (Mower 2015: 1–2).[4] Implicit bias is presented as a challenge to current approaches.

Project Implicit provides a website where individuals can take Implicit Association Tests (IAT) to assess their implicit biases (Project

Implicit). The program assesses how quickly positive and negative traits are associated with certain categories of people; a delayed response in associating a particular group with a positive trait serves as evidence of implicit bias. Consider the case of implicit bias with regard to the races of European Americans and African Americans.[5] While completing the Race IAT, if you correlate members of some races with positive traits more quickly than you do other races, then at an implicit level the positive correlation is not as strong or immediate for the members of races to which you had a delayed response. As described at the end of the Race IAT, "The interpretation is described as 'automatic preference for European American' if you responded faster when *European American* faces and *Good* words were classified with the same key than when *African American* faces and *Good* words were classified with the same key" (Project Implicit). The magnitude of your result is also described as slight, moderate, strong, or little to no preference (Project Implicit). Of course, one could also rank as having an automatic preference for African Americans at the end of the test.

The results of Project Implicit are relevant to moral theory, practice, and pedagogy. Alarmingly, 75 percent of all people have an automatic "White preference" in which a stronger association occurs between white people and positive traits (Banaji and Greenwald 2013: 47; also quoted in Mower 2015: 4). Although an automatic preference should not be equated with racism insofar as racism involves intentional acts of discrimination characterised by hostility, rage, and disrespect, it is still disconcerting that despite one's "consciously held, rationally chosen, reflective and principled beliefs, one's own mind and the categories and concepts therein might be rigging the game" (Mower 2015: 5). Even if automatic preferences are not the equivalent of racism, they nonetheless predict discriminatory behavior (Mower 2015: 5). The results of Race IATs "are a stronger predictor of discriminatory behavior than self-reports by subjects with overtly racist attitudes" (Banaji and Greenwald 2013: 52; also quoted in Mower 2015: 6).

Education, as currently practiced, does not serve as an adequate corrective to such ignorance. Although on self-report measures the more educated a person is, the weaker the association between "Black" and "weapons," when the same people were given an IAT their educational level had no effect on implicit beliefs and associations (Mower 2015: 6). One could therefore argue that educational endeavors as currently practiced fail, at least in some domains, to have an impact on ethically problematic implicit associations and the resultant discrimination that occurs (Mower 2015: 7).

Debiasing Strategies

In this section, I offer a number of teaching and learning strategies to help work against the operations of explicit and implicit bias. Corrective measures for the suite of biases outlined by Horton, and the empathetic biases addressed by Hoffman, are explored. These include such practices as enhanced thought experiments, interviews, research projects, in-depth role-playing, and appropriately morally deferential experiential service-learning working directly with oppressed populations, including through action projects. Moral imagination can be stretched through adopting differing moral lenses and engaging in exercises that encourage multiple empathizing. Ways in which moral imagination can be grown through the use of art, and literary narrative in particular, are identified. Relatedly, games and videos as creative learning devices have been shown to reduce harmful bias. Critical scrutiny focused on personal biases (including teacher bias), and developing curriculum focused on moral literacy, likewise contribute to working against the problematic operations of bias. Such moves of necessity span from individual to public action given the environmental components of the operations of bias. As for the operations of implicit bias, Mower recommends fostering a culture of sensitivity in classrooms, committing ourselves to altering patterns of bias, intentional selection of association formation, and collaborative critical analysis. Thus, the opportunity to shape ourselves through intentional environment construction and avoidance of undesirable environments surfaces as an important ethical technique. Finally, the potential contribution of loving-kindness meditation is addressed.

Horton refers to debiasing strategies as those that are intended to counteract the influence of bias (Horton 2004: 551). Fortunately, for philosophers, they are, for the most part, a standard part of philosophical practice: generating counter-arguments, articulating reasons in support of opposing beliefs, and expressing the line of thinking that leads one to hold particular beliefs (Horton 2004: 552). Part of the benefit of such practices is that, when performed with regularity, they can serve to inculcate habits of thinking and perceiving. Insofar as bias operates with automaticity, it strikes me as reasonable to imagine an adequate counter to such bias needs to be habituated to the point where it is automatic. Techniques are recommended such as imaginatively taking the place of the person or people that the actions discussed will impact; this encourages adopting perspectives at a remove from one's own (Horton 2004: 553). Insofar as one can successfully adopt a perspective other than one's own, belief perseverance bias can be challenged. If one adopts an alternative perspective, one can no longer uncritically gravitate toward information that confirms one's existing beliefs. One is forced to imaginatively take another perspective. John

Rawls's thought experiment in *A Theory of Justice*, in which he employs the technique of adopting a veil of ignorance in terms of one's placement in society, serves as an illustration. One is asked to consider what kind of world they would construct if they did not know what social position, geographical location, and so on, they would take on in that world. After engaging with this thought experiment, students tend to recognize that more of their life is a matter of luck and less a matter of what is deserved.

One worry about thought experiments is their level of abstraction, particularly given the import of context for moral decision-making (Urban Walker 1997). There is also a concern about the arrogance involved in assuming one can imaginatively take the place of another by a simple act of ratiocination. Ways to counteract such harms might include having students, after they have completed a thought experiment, conduct research on related issues, and/or interview people who face related dilemmas in real life, and/or invite guest speakers who work in a related field and/or are members of oppressed groups that are discussed in the thought experiment. Such methods can serve to enhance traditional thought experiments.

An additional concern that is offered with regard to current methods of teaching ethics is that emphasizing moral reasoning skills fails to adequately reflect how moral decision-making happens. The claim here is that emphasizing reasoning skills as "teaching and learning goals in ethics relies on the naïve, pre-psychological notion that moral reasoning processes typically precede the formulation of a moral judgment, which in turn becomes the basis of action or choice" (Maxwell 2016: 81). Multiple theorists cite evidence illustrating that it is not in fact the case that moral reasoning processes typically precede the formulation of a moral judgment (Railton 2014: 827–29). Robert Zajonc (1980) introduced the idea of "affective primacy" wherein new sensory input engages our affective system almost immediately (Railton 2014: 827). Some theorists forward a dual process model of the mind wherein the affective system (System 1) is the substrate for a distinctive, automatic mode of processing, while the higher cortical regions serve as the substrate for more controlled and effortful modes of functioning (System 2). (Railton 2014: 827). For Jonathan Haidt, rational process primarily functions as an afterthought attempting to justify what has already been decided on intuitive grounds; it is only very rarely that reason is able to exercise a small measure of control over what is done (Haidt 2012: 46; also quoted in Johnson 2014: 46; Haidt 2001: 823, 830). Such views are contested, to be sure, but if such is the case, one can argue that focusing on traditional approaches offers little toward shaping lived moral responses to actual situations. However, even with a dual process model of the mind that assumes

the primacy of System 1 for moral assessment, there is still a case to made for teaching critical thinking to help work against bias. Bruce Maxwell suggests we take a closer look at what precisely teaching students moral reasoning is meant to do. Maxwell encourages imagining models of ethical deliberation as frameworks as opposed to recipes. If one embraces this reconceptualization, one can remain consistent with new developments in moral psychology while still recognizing the merits of teaching reasoning skills (Maxwell 2016: 82). Models "of moral deliberation can serve as an analytic framework for breaking an impasse of uncertainty about an ethical problem—rather than as a kind of algorithm which, when applied to an ethical problem, produces a reliable moral judgment on which to base the ethically right course of action in the circumstances" (Maxwell 2016: 82). Critical thinking therefore is not meant as a substitute for System 1; rather, it can be used to put the responses of System 1 into critical conversation with reflected values. Such considerations flag the importance of generating social structures that create and habituate the conditions for ethically and epistemically appropriate emotional responses.

Another potentially promising approach would be to conduct in-depth role-playing. Mower, for example, has led classes in which students adopt a heuristic that includes identifying moral questions, engaging with a script, adopting perspectives, examining what is found, and coming up with moral answers (Mower 2008: 2). Students complete a variety of role-playing exercises that enhance emotional literacy (emotional recognition in self and other), contextual literacy (imagining morally charged scenarios vividly and realistically), and moral literacy (in which first- and third-person perspectives can be imagined) (Mower 2008: 3). Role-playing invites a more visceral, emotional, and tangible method for imagining the experience of an other. Student learning experience need not be limited to the classroom, however.

When students engage directly with oppressed populations through experiential learning, the biased epistemic conformity that arises from particular locations can be challenged in a powerful way. If the encounters afford vivid, concrete examples, an automatic bias for tending to that which is vivid and concrete could be put to good use. We may find that, insofar as here and now bias is operative, it can be strategically used to our advantage. In the absence of being able to mitigate here and now bias, one could argue it is worthwhile to work with it so that it better serves moral goals. The moral understanding of oppressed others in our local communities can be increased due to the immediacy and proximity of the moral harms. Here and now bias might serve to support positive moral action when students, through action projects and service-learning, engage with populations who directly experience the harms of moral injustice and lack of care. The

same can happen when our oppressed students make vivid the concrete experiences of lived oppression when they generously share that experience in classroom discussion. The selective evidence gathering that wishful-thinking bias encourages could also be challenged. If one engages head-on with the nuances of an actual issue and the people it affects, epistemic support for choosing to ignore huge parts of the story is lessened. Service-learning, action projects, and interviews are recommended to help ameliorate bias. These are approaches in which dialogue is opened and experiential understanding is thereby opened.

However, students must be prepared for such encounters. Ensuring students are prepared to listen, and maybe even listen so well that they give up epistemic authority for a time, will be important to ensure complexity is foregrounded and the simplifications that bias invites are called into question. Without adequate preparation, students might extend their confirmation bias to their service-learning experience or action project. Also of concern is the inappropriate "ah ha" moment. It is important to avoid a superficial engagement in which students come to think they understand the other via an "ah ha" moment through reading a text or seeing a film about a person who is differently placed (Boler 1999: 157). So too in the case of interacting directly with differently placed others. A deep understanding of the perspective of the other requires a commitment to meaningful engagement that reflects the experience of the other as articulated by the other, and a humble recognition of the limits and lacunas in one's own perceptual frames. Recognition that we all adopt moral lenses through which we see issues is therefore relevant.

Researching how different moral lenses differently configure moral issues could also be beneficial, as could practices in which students exercise their moral imaginations specifically to increase empathy. As Susan Sherwin argues convincingly, each moral lens reveals some aspects of moral issues while concealing others (Sherwin 1999). It is only by trying on multiple lenses that one can see moral issues more fully. The insights derived from multiple perspectives generate information needed for an informed perspective. Hoffman recommends adopting exercises that encourage "multiple empathizing" wherein individuals imagine what it would feel like for those harmed by a particular action (Hoffman 2000: 297). Through developing students' skills with regard to moral imagination, and by vesting experiential authority in those who actually experience such harms, students could work against empathetic biases such as familiarity bias, in-group bias, friendship bias, and similarity bias. Students who do not already experience intersectional oppression could engage in activities that invite moral imagination in the direction of what it might feel like to have a matrix of oppressive institutional limitations and harms constantly at work in

one's life. Such a move would have to be inherently respectful and involve moral deference toward the experience of oppressed populations (Thomas 1992–1993). This, again, would help speak to the need for more vivid engagements with issues to instigate motivation for action (Horton 2004: 553).

Another means by which bias can be counteracted is through surveying one's own thinking and taking a critical stance toward one's own position to look for evidence of biases; if one finds such evidence, then the issue must be re-examined (Horton 2004: 554). Such reflection is required because bias, both explicit and implicit, still occurs among students on college campuses (Boyson and Vogel 2009: 16). Indeed, teachers themselves must also be vigilant, not only in terms of the literature presented to students, but also in their interactions with students, to take a step back and look for evidence of bias. One way in which teachers can be biased is through the halo effect. The halo effect is a type of confirmation bias wherein a positive interpretation of a person in a particular domain causes one to generate a positive interpretation of that person in another domain, even though the jump is unsubstantiated. It refers to the tendency to perceive a person's positive or negative traits as extending from one area of their personality to another (Wilke and Mata 2012: 532). The halo effect can indicate times in which liking an individual is confused with objective judgments; for example, a professor's grade for a student being elevated due to liking the student (Wilson and Brekke 1994: 129; Nisbett and Wilson 1977: 244). One way to counter this sort of bias is to mark assignments without awareness of the name of the student who submitted the work; for example, requesting that they use their student number instead of their name on submitted assignments so as to maintain anonymity.

Another fruitful method for countering bias in the classroom is to develop curriculum clearly grounded in, and oriented to, moral literacy. Nancy Tuana gives a comprehensive analysis of the fundamental elements of moral literacy. In Tuana's view, moral literacy involves ethics sensitivity, ethical reasoning skills, and moral imagination. Ethics sensitivity involves "(1) the ability to determine whether or not a situation involves ethical issues; (2) awareness of the moral intensity of the ethical situation; and (3) the ability to identify the moral virtues or values underlying an ethical situation. These abilities are complex and require training and practice to master" (Tuana 2007: 366). Ethical reasoning skills require "(1) an understanding of the various ethical frameworks; (2) the ability to identify and assess the validity of facts relevant to the ethical situation, as well assessing any inferences from such facts; and (3) the ability to identify and assess the values that an individual or group holds to be relevant to the ethical issue under consideration" (Tuana 2007: 369). The final component, moral imagination,

is defined by Mark Johnson as the "ability to imaginatively discern various possibilities for acting in a given situation and to envision the potential help and harm that are likely to result from a given situation" (Johnson 1993: 202; also quoted in Tuana 2007: 374). Moral imagination includes abilities such as: empathy for the feelings of others, the ability to imagine ourselves in the situation of someone else, creatively thinking outside the box, attunement to the complexities of a situation, developing trust and acting in helpful ways, appreciating the feelings of others even when they differ significantly from our own, sensitivity to nonverbal cues, personal ownership, and habituation of ethical behavior—including a felt sense of responsibility for our actions, as well as desire for both justice and virtuous habits (Tuana 2007: 375).

A number of the methodologies recommended for developing adequate moral literacy simultaneously work against the operations of bias. In particular, the operations of moral imagination overlap in key ways with the methods recommended above for countering biases. To determine what constitutes an ethical situation and to correctly gauge its intensity and the moral values at play, one needs the ability to see through multiple lenses beyond one's own biased lens. Ethical reasoning skills, developed with an eye toward minimizing bias, can correctly identify multiple perspectives as well as what follows from them. Moral imagination is an ability that can be developed and practiced skillfully so as to directly counter the operations of bias. Moral imagination's ties to empathy, imaginatively taking the perspective of another, and appreciating the feelings of others all point to methods for questioning the limitations and background assumptions of one's own position.

Thus, what moral imagination amounts to and how it can be nurtured merit further exploration. Let us return to Mark Johnson's account. He argues that moral understanding is in a large part imaginatively structured through concepts "with prototype structure, semantic frames, conceptual metaphors, and narratives" (Johnson 1993: 198). As such, we have two tasks if we are to be morally insightful and sensitive; first, we "must have knowledge of the imaginative nature of human conceptual systems and reasoning," and second, we "must cultivate moral imagination by sharpening our powers of discrimination, exercising our capacity for envisioning new possibilities, and imaginatively tracing out the implications of our metaphors, prototypes, and narratives" (Johnson 1993: 198). Moral development becomes less about learning how to apply moral laws and more about "refining our perception of character and situations and of developing empathetic imagination to take up the part of others" (1993: 199). On Johnson's analysis, we need to cultivate "an ability to imaginatively discern various possibilities for acting in a given situation and to envision the potential help and harm that are likely to result from a given action" (1993: 202). Affective

and rational dimensions are activated in that you both need to be able to deduce likely consequences as well as the relevant affective states. Through questioning currently dominant moral prototypes, semantic frames, conceptual metaphors, and narratives, we are enabled to think differently and to imagine newly. Moral imagination can be used to break through problematic sedimented modes of thought and feeling that become engrained and even invisible to ourselves if unchallenged.

Methods for cultivating moral imagination include engaging with various forms of art. For example, Dan Johnson, Brandie Huffman, and Danny Jasper found that reading narrative fiction successfully changed individuals' race boundaries by reducing categorical race perception and inhibiting the tendency to categorize mixed-race individuals with angry expressions as out-group members (Johnson, Huffman, and Jasper 2014: 88). The narrative presented was taken from Shaila Abdullah's novel *Saffron Dreams* (2009: 57–64) and involved a scenario in which the protagonist, Arissa, who is a strong-willed Muslim woman, is assaulted due to her ethnicity and exhibits remarkable courage by standing up to her attackers' ethnic and religious slurs (Johnson, Huffman, and Jasper 2014: 84). Reading this narrative fiction ameliorated "biased categorical and emotional perception of mixed-race individuals" (Johnson, Huffman, and Jasper 2014: 88).

Johnson et al. hypothesize that their findings have implications for the perception-action coupling model of empathy, wherein those who experience empathy for another will exhibit brain activity in the areas involved in performing action X when they perceive another person performing action X (Johnson, Huffman, and Jasper 2014: 88). Narrative fiction, it is hypothesized, might help traverse the empathy gap which indicates less perception-action coupling for out-group members (Johnson, Huffman, and Jasper 2014: 88). Support for this notion can be found in the "growing evidence that reading a story engages many of the same neural networks involved in empathy and developing a theory of mind of others" (Mar 2011; Speer, Reynolds, Swallow, and Zacks 2009; also quoted in Johnson, Huffman, and Jasper 2014). More in-depth study is needed, as well as longitudinal study, particularly given the findings of Calvin Lai et al. (2016). In a study seeking to test implicit racial preferences across time, they used a vivid second-person story that was emotionally involving and provided a counter-stereotypic scenario (Lai et al. 2016: 1004–05). This approach was the most effective short-term intervention for shifting implicit preferences; however, for all nine interventions tested, the effects lasted less than a couple of days (Lai et al. 2016: 1011,1014).[6] The findings provide support for pursuing methods that have been shown beneficial for long-term change in implicit preferences and stereotypes; namely, prolonged everyday experience to interventions (Lai et al. 2016: 1013). Examples

include taking a class with a female professor (Dasgupta and Asgari 2004; Stout, Dasgupta, Hunsinger, and McManus 2011), and taking a semester-long class on the topic of prejudice and intergroup conflict (Rudman, Ashmore, and Gary 2001) (Lai et al. 2016: 1013). Habituating empathetic perspective-taking through consistently utilizing narratives in class that seek to enable such skills may therefore be an effective method for reducing bias.

Focusing on emotional components of moral imagination might likewise be useful. When instructed to be empathetic when reading about the discrimination of black people, the difference between participants' evaluations of black and white Americans was eliminated (Stephan and Findlay 1999; also quoted in Paluck and Green 2009: 348). Thus, narrative engagement seems to provide one worthwhile venue for bias reduction that could be included in philosophical classes with relative ease. Indeed, many of us already use creative narrative fiction in the classroom. Questions that remain include the following: precisely how does narrative function to ameliorate various forms of bias, and what texts provide uniquely adept narratives for ameliorating various forms of bias? More broadly, art can encourage moral imagination through challenging complacent assumptions about matters of fact and how problems and solutions should be framed (Elgin 2002: 12). Art can be used to push boundaries, reconfigure domains, and illuminate unusual perspectives; it can help to disrupt, disorient, and force one to look beyond the status quo (Elgin 2002: 12).

A further interesting finding for lessening bias pertains to the use of videos and games. Carey Morewedge et al. found that utilizing a video or games can have significant debiasing effects across a variety of contexts affected by the same bias (Morewedge et al. 2015: 137). The video scenario involves participants watching a training video wherein heuristics are defined, and ways in which heuristics can lead to incorrect inferences are flagged (Morewedge et al. 2015: 133). Then multiple biases are defined (bias blind spot, confirmation bias, and fundamental attribution error), vignettes are used to illustrate individuals committing such biases, and mitigating strategies are suggested (Morewedge et al. 2015: 133). In the game scenario, participants played a game designed to elicit and mitigate the forms of bias mentioned above (Morewedge et al. 2015: 133). It involves taking a first-person point of view; players make judgments that test the degree to which they exhibit biases, and experts explain each bias and also give narrative examples at the end of each level of the game (Morewedge et al. 2015: 133). Participants are then "given personalized feedback on the degree of bias they exhibited. Finally, participants perform additional practice judgments of confirmation bias (five in total) and receive immediate feedback before the next level begins or the game ends"

(Morewedge et al. 2015: 133). The results suggest that even a single training intervention "can have significant debiasing effects that persist across a variety of contexts affected by the same bias" (Morewedge et al. 2015: 137). Notably, participants in the game scenario experienced large reductions in cognitive biases immediately, and the reductions persisted at least two or three months later (Morewedge et al. 2015: 137). The games had greater efficacy than the videos, and Morewedge et al. hypothesized that this is due to the impact of personal feedback and practice which operate to increase the debiasing effects of the training (Morewedge et al. 2015: 137). In the game, the player searches for a missing neighbor and exonerates her from criminal activity (Morewedge et al. 2015: 133). Given Johnson, Huffman, and Jasper's (2014) findings, one could suggest that the narrative element of the game coupled with the first-person perspective could be contributing factors to the game's success in reducing biases. Whatever approaches are adopted, evidence of their efficacy is needed.

Elizabeth Paluck and Donald Green (2009) ask the empirical question of what works when it comes to prejudice reduction. After reviewing the existing data they concluded that the causal effects of numerous widespread prejudice-reduction interventions remain unknown (Paluck and Green 2009: 339). But they also state that cooperative learning is a viable method for reducing prejudice and that media, reading, and other forms of narrative and normative communication are successful techniques in the real world (Paluck and Green 2009: 358). Such tools merit further study given their potential for reducing prejudice "through narrative persuasion, social norms, empathy, perspective taking, and extended contact" (Paluck and Green 2009: 358). Paluck and Green also believe focusing on the persuasive and positive influence of peers is a promising area of study for prejudice reduction (Paluck and Green 2009: 358). Class discussions could therefore take on new dimensions aimed at prejudice reduction through the narratives shared and by thoughtfully tending to how students are influencing, and influenced by, their peers.

Given wider cultural, social, political, and economic means for engraining particular biases, focus on the classroom alone will be fundamentally incomplete if our goal is to minimize bias. For example, we are under a constant barrage of carefully framed advertising designed to have specified impacts, such as convincing us that we are not good enough as we are, and that to be worthy we ought to move closer to the mythic ideal through product consumption (Lorde 1984). The news is carefully framed either to support or shift politics, our sense of the state of the world, and what in it is worthy of focus (war, murder, death, the economy, etc.) and what is not (namely the many good-hearted actions that happen daily the world over). Matthieu Ricard points to the banality

of the good—our society hangs together only through remarkable acts of compassion and generosity and care happening daily, but it happens so often that they fade into the background (Ricard 2015: 95–102). It certainly does not tend to make the news, and our views about our fellow humans are shaped accordingly. When the news focuses on covering particular genders, sexualities, races, embodiments, and so on, in stereotype-reifying ways bias is thereby supported. Moreover, one's own cultural lens is further entrenched when one is surrounded by others who participate in that culture. Thus, combating bias cannot stop at the individual level. Investing in social, political, and economic moves that work against human tendencies toward problematic bias is therefore an important goal for action and growth.

Students who work on action projects and service-learning can effectively make explicit ways in which acting on more than an individual level is possible. Through action projects, students can engage their wider school, city, state, nation, and global communities. When designing assignments, part of the goal can be to highlight the ways cultural changes have been brought about in the past, and the ways students today can choose to use their energies effectively for activating that change.[7] Thus, anti-bias work can reach out beyond individual students in hope-inducing ways. Through action projects, they can be supported in their efforts to change the design of social, economic, and political systems through political engagement.

One could argue, however, that none of the above strategies adequately deal with currently operating implicit biases. These chronic biases seem to operate beyond the reach of explicit thought and are seemingly immovable. Mower, however, suggests change is indeed possible, and she provides four methods for countering the operations of implicit bias in the classroom. First, we ought to foster a culture of sensitivity in our classrooms that promotes actively examining patterns of thought and looking for evidence of bias (Mower 2015: 12). Second, we must commit ourselves to altering the patterns of implicit bias we discover (Mower 2015: 12). If the news insists on presenting people of colour in a demeaning and violent light, for example, then students and teachers can actively seek out stories, articles, and pictures of people of colour in a respectful and peaceful light.[8] Third, we must commit ourselves to the selection of more morally acceptable associations so that the very way we are formed is the result of moral choice rather than merely a reflection of unchosen current and contingent social practices (Mower 2015: 13). Not only must there be a commitment to a principled selection of the associations to which we repeatedly expose ourselves, we must also inculcate the morally preferable associations (Mower 2015: 13). Fourth, this work cannot be a solitary endeavor; there must be a collaborative critical analysis

wherein group discussions, projects, and activities can benefit from group diversity (Mower 2015: 13).

We have the opportunity to shape ourselves, and to shape ourselves differently. For example, if we binge watch television that confirms stereotypes, we internalize those stereotypes. If we spend most of our time with close-minded and judgmental people, we stay closed-minded and judgmental. Although not unlimited, there are a number of choices we can make that will shape us in new directions. For example, choosing to stop watching television can have a remarkable impact on how one frames the world and its members. Granted, one need not necessarily go this far, but being as intentional as possible with regard to what we spend our time doing—what we nurture our minds, hearts, and imaginations with—will impact not just explicit beliefs and values, but the implicit operations of bias as well. If certain ideas do not gain a foothold in the first place, then the worry of implicit bias either need not arise at all or not arise in such a pressing way. In keeping with the insights of a positive psychology approach, it would be helpful to identify individuals with the least explicit and implicit biases and find out what is consistent among that population.

Additional critical attention should focus on how environments shape bias or the lack thereof. A sole focus on heuristics and biases in relation to human cognitive systems—cognitive systems conceptualized as being largely separated from complex wider environments and systems—fails to provide a comprehensive picture of how bias functions (Wilke and Mata 2012: 531–32). An emphasis is needed on the role of environment in what constitutes adaptive behavior (Wilke and Mata 2012: 532). Part of the issue with cognitive bias is that the human cognitive system "was designed to make inferences about the external world based on imperfect cues that could lead to errors in some situations" (Wilke and Mata 2012: 531). But such failures cannot be adequately observed in relative isolation, for "the success of reasoning strategies depends on the structure of the environment" (Wilke and Mata 2012: 532). Thus, the design and impact of institutional, social, political, economic, and other related structures requires deep analysis regarding the biases they perpetuate. We neglect addressing problematic institutions at our peril, for those institutions shape our biases (Jacobson 2016: 184). However, we cannot focus solely on institutional change; we neglect addressing problematic attitudes at our peril as well, given that attitude change can help fuel the desire for institutional change (Jacobson 2016: 184).

Relatedly, Kenyon and Beaulac (2014) argue in favor of the need for new approaches to teaching about bias when teaching critical thinking. They recommend methods which include outcomes-oriented and externally-mediated approaches that are not limited to individual-

ist and internalist character-driven approaches (Kenyon and Beaulac 2014: 352). Kenyon argues that a strategy for general debiasing is multi-leveled, and includes things like: extensive targeted training, inculcating debiasing habits, imparting social attitudes that tolerate and respect conciliating and changing one's mind, and the creation of reasoning infrastructure (Kenyon 2014: 2545). Moreover, he contends that the projects of acquiring and imparting effective reasoning and communication skills are "profoundly socialized projects. There is no realistic hope of cultivating the relevant kinds of reliable belief-forming processes without a vast network of background educational support structures, and probably not without a great deal of ongoing social and institutional reasoning support as well" (Kenyon 2014: 2545–46). In a similar vein, Natalia Washington and Daniel Kelly flag the importance of looking at the interaction of individual and collective responsibility, and they recommend attending to the ways institutions can intentionally structure environments so as to work against bias and prejudice (Washington and Kelly 2016: 27). Setting up environments such that bias does not arise in the first place, or the likelihood of it arising is minimized, is a useful approach for coping with bias. Training students in the intentional subjection of oneself to various debiasing aids in the form of "nudges, infrastructure and institutions" in advance of the circumstances of bias is a novel approach to bias reduction (Kenyon and Beaulac 2014: 356). The need for intentionally shaping one's environment so as to manifest desired behavior is flagged here. One could argue such skills are needed in philosophical practice more generally, including with regard to teaching ethics.

To work most effectively against bias, the operations of both implicit and explicit bias need to be addressed. Their operation at the individual level and various group levels, and the ways in which bias overlaps, supplements, or grows with other epistemic and moral deficiencies must also be explored. The merits of focusing on how wider systems shape individual responses (including biased responses) include an appropriate concern with systems of power and dominance and the need for large-scale change. If this is done at the expense of individual responsibility for working against bias, however, then options that are potentially far more realizable in a shorter amount time and provide immediate correctives and empowerment are ignored. Thus, strategies that attend the need for both individual and group action are simultaneously required. Although only a small portion of concerns arising from the intersection of ethics, pedagogy, and bias were addressed, I hope to have made a strong case for further research in this vein. The fact that education served to decrease explicit bias remains heartening. Even though implicit bias was still operational in the Race IAT, and even though there was a delay in terms of categorizing a particular race

with a particular positive quality, insofar as the explicit categorization was the correct one, then education has been moving in the right direction at least at the explicit level. The next step is to discover how to move far more quickly and more effectively through tending to the insights of moral psychology with regard to the operations of bias. A number of methods were discussed above. I conclude by turning to an exciting direction for resisting bias that has not yet been richly explored in ethical pedagogy; namely, the use of mindfulness practice in the classroom.

There is a burgeoning area of literature in mindfulness studies that provides a potential wealth of information for philosophical pedagogy. Engaging in ten minutes of mindfulness meditation has been found to (1) reduce implicit age and race bias due to a weakening of automatic associations (Lueke and Gibson 2014), and (2) reduce discrimination (Lueke and Gibson 2016). Participating in a loving-kindness meditation practice during a six-week period significantly decreased implicit bias toward Black people and homeless people, while simply discussing loving-kindness in a similar group setting did not (Kang, Dovidio, and Gray 2013: 5). Alternatives to loving-kindness meditation, for example relaxation exercises, are not likely to have the same robust and generalized effect for bias reduction (Kang, Dovidio, and Gray 2013: 6). I now turn to a specific study that suggests an approach that can be utilized in the classroom to decrease bias.

Stell and Farsides conducted a study on the relative effects of other-regarding and non-other-regarding positive emotions on bias (Stell and Farsides 2016: 140). In particular, the Buddhist practice of loving-kindness meditation was used, which "seeks to self-regulate an affective state of unconditional kindness towards the self and others" (Stell and Farsides 2016: 140–41). Although it can be practiced in a variety of ways, repeating phrases such as "may you be happy and healthy" while visualizing a person experiencing the outcome of this well-wishing is a common approach (Stell and Farsides 2016: 141). Loving-kindness meditation has been shown to increase positive affect as measured by self-report and autonomic processes such as vagal tone, and it has also been shown to significantly increase self-reported and implicit positive affect toward a photograph of a neutral stranger (Stell and Farsides 2016: 141).[9] Short, seven-minute exposures to loving-kindness meditation directed toward a member of a racial out-group decreased racial bias towards that out-group as measured by the IAT (Stell and Farsides 2016: 145). Racial bias decreased by both increasing controlled processing and decreasing automatic processing (Stell and Farsides 2016: 145). The improved appraisals were not limited to the individual (the photograph) on whom the loving-kindness meditation was focused; it extended toward those who were categorically

related through race to the focused-on individual (Stell and Farsides 2016: 145).[10] Other-regarding emotion generated by loving-kindness meditation was a predictor of bias reduction, but non-other-regarding positive emotions did not provide evidence of bias reduction—thus, the social locus of the emotion matters (Stell and Farsides 2016: 146). Importantly, for our purposes, seven minutes of loving-kindness meditation can reduce bias against members of a racially oppressed group that is about to be discussed in class. Further research might show that loving-kindness meditation reduces bias not just in cases of racial bias but other forms as well. The habituation of practices that result in self-reported and autonomic evidence of decreased bias can help create the conditions for more ethical ways of being. The toolbox for countering bias may increase in size and content as additional findings are revealed with regard to mindfulness training as it intersects with bias. Efficacy might reasonably be linked to the regularity of engagement with such practices, given that regular practice supports their habituation. Clea Rees suggests that habituation of egalitarian commitments is crucial to strategies of resistance regarding automatic processing; however, the sort of community one is embedded in likewise needs tending to so that the community itself is committed to resisting implicit bias (Rees 2016: 192–93). Cognitive systems automatically adapt our responses to social environments; as such, community support for egalitarianism is essential for developing and maintaining egalitarian virtues (Rees 2016: 210).

Conclusion

The intersection of ethics, bias, and pedagogy is an area ripe for further work. I have explored a variety of forms of bias, which included belief perseverance through ignorance; epistemic conformity due to social pressure; the disproportionate weight given to vivid, concrete, examples over abstract ones; wishful thinking and the associated less-rigorous assessment of information that fails to support the desired conclusion; empathetic biases such as in-group, friendship, similarity, and here and now bias; and differential and unjustified attributions of rhetorical credibility. Debiasing strategies were looked at such as utilizing enhanced thought experiments, experiential learning, role-playing, research projects, service-learning, action projects, interviews, adopting multiple moral lenses, critical scrutiny focused on personal biases and engaging in corrective practices, and developing curriculum focused on moral literacy, which includes expanding the horizons of moral imagination. Individual and public action are addressed, given that bias operates in particular contexts and is informed by our social, political, economic, and ecological environments. A number of methods

for correcting the operations of implicit bias were addressed, including creating a culture of sensitivity in the classroom, commitment to altering patterns of bias, intentional selection of association formation, collaborative critical analysis, use of narrative and the arts, intentional environmental creation and avoidance of particular environments so that the bias is not habituated in the first place, and recent findings regarding the relationship between loving-kindness meditation and explicit and implicit bias reduction. Until problems are identified, solutions will not be forthcoming. Exploring the operations of implicit and explicit bias, as well as attendant issues, is a growing area of research. As such, the problems bias poses for ethics generally, and for ethics pedagogy in particular, is an exciting area for further development.

Notes

1. Although I cannot discuss all forms of bias in detail due to space constraints, and the forms referred to in what follows are just scratching the surface, it is worth mentioning diverse operations of bias so as to highlight areas needing further study.

2. The stereotyping of, and attendant operations of bias toward, groups of individuals in this way is not limited to fellow humans. Examples of bias against non-humans include speciesism (Singer 2002) and naturism (Warren 2005).

3. This paper is, in part, a response to that call. However, I focus on bias more widely, in both its explicit and implicit manifestations.

4. For additional reasons to be concerned, see Kretz 2012 and Kretz 2014.

5. By discussing these categories, I am adopting the terminology used for the Race IAT. Although tending to racial categories is important insofar as they currently operate to entrench social, political, and economic forms of oppression, saying that is not to adhere to the view that race is a viable stand-alone ontological qualifier removed from contexts of oppression.

6. The nine interventions included vivid counterstereotypic scenario, practicing an IAT with counterstereotypical exemplars, shifting group boundaries through competition, shifting group affiliations under threat, priming multiculturalism, evaluative conditioning, evaluative conditioning with the GNAT, using implementation intentions, and faking the IAT (Lai et al. 2016: 1004–06).

7. For service-learning and action projects it is important to check your institutional policy for such activities and to identify requirements for waivers and the like.

8. To overcome the effects of implicit bias on cognitive processing, Clea Rees recommends imagining counter-stereotypic exemplars or imaging oneself in another's shoes (Corcoran, Hundhammer, and Mussweiler 2009; Dasgupta and Greenwald 2001; Dasgupta and Asgari 2004; Blair, Ma, and Lenton 2001; and Galinsky and Moskowitz 2000 [also quoted in Rees 2016: 196]).

9. Although some forms of positive affect have been shown to increase prejudiced thoughts and feelings toward racial out-groups (Huntsinger et al. 2009; also quoted in Stell and Farsides 2016: 140), other forms such as "awe and nurturant love contraindicate positive emotions' usual reliance on automatic decision-making" (Griskevicius et al. 2010; also quoted in Stell and Farsides 2016: 140). Loving-kindness meditation is another example in which positive emotion results in decreased bias.

10. Although priming negative construals of the in-group have also been shown to reduce bias, such as priming despised white figures such as Adolf Hitler and Ted Bundy, loving-kindness meditation serves as a more acceptable general method of reducing bias (Lai et al. 2014; also quoted in Stell and Farsides 2016: 145).

Bibliography

Abdullah, Shaila. 2009. *Saffron Dreams*. Ann Arbor, Mich.: Modern History Press.

Blair, Irene V., Jennifer E. Ma, and Alison P. Lenton. 2001. "Imagining Stereotypes Away: The Moderation of Implicit Stereotypes through Mental Imagery," *Journal of Personality and Social Psychology* 81(5): 828–41. https://doi.org/10.1037/0022-3514.81.5.828

Boler, Megan. 1999. *Feeling Power: Emotions and Education*. New York: Routledge.

Boyson, Guy A., and David L. Vogel. 2009. "Bias in the Classroom: Types, Frequencies, and Responses," *Teaching of Psychology* 36(1): 12–17. https://doi.org/10.1080/00986280802529038

Corcoran, Katja, Tanja Hundhammer, and Thomas Mussweiler. 2009. "A Tool for Thought! When Comparative Thinking Reduces Stereotyping Effects," *Journal of Experimental Social Psychology* 45(4): 1008–11. https://doi.org/10.1016/j.jesp.2009.04.015

Dasgupta, Nilanjana, and Shaki Asgari. 2004. "Seeing Is Believing: Exposure to Counterstereotypic Women Leaders and Its Effect on the Malleability of Automatic Gender Stereotyping," *Journal of Experimental Social Psychology* 40: 642–58.

Dasgupta, Nilanjana,a nd Anthony G. Greenwald. 2001. "On the Malleability of Automatic Attitudes: Combating Automatic Prejudice with Images of Admired and Disliked Individuals," *Journal of Personality and Social Psychology* 81(5): 800–14. https://doi.org/10.1037/0022-3514.81.5.800

Elgin, Catherine. 2002. "Art in the Advancement of Understanding," *American Philosophical Quarterly* 39: 1–12.

Galinsky, Adam D., and Gordon B. Moskowitz. "Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-Group Favoritism," *Journal of Personality and Social Psychology* 78(4): 708–24. https://doi.org/10.1037//0022-3514.78.4.708

Govier, Trudy. 1993. "When Logic Meets Politics: Testimony, Distrust, and Rhetorical Disadvantage," *Informal Logic* 15(2): 93–104. https://doi.org/10.22329/il.v15i2.2476

Griskevicius, V., M. Shiota, and S. Neufeld. 2010. "Influence of Different Positive Emotions on Persuasion Processing: A Functional Evolutionary Approach," *Emotion* 10(2): 190–206. https://doi.org/10.1037/a0018421

Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108(4): 814–34. https://doi.org/10.1037/0033-295X.108.4.814

Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York: Vintage Books.

Hoffman, Martin. 2000. *Empathy and Moral Development: Implications for Caring and Justice*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511805851

Horton, Keith. 2004. "Aid and Bias," *Inquiry: An Interdisciplinary Journal of Philosophy and the Social Sciences* 47: 545–61.

Huntsinger, J. R., S. Sinclair, and G. L. Clore. 2009. "Affective Regulation of Implicitly Measured Stereotypes and Attitudes: Automatic and Controlled Processes," *Journal of Experimental Social Psychology* 45(3): 560–66. https://doi.org/10.1016/j.jesp.2009.01. 007

Jacobson, Anne. 2016. "Reducing Racial Bias: Attitudinal and Institutional Change," in *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, ed. Michael Brownstein and Jennifer Saul. Oxford: Oxford University Press.

Johnson, Dan R., Brandie L. Huffman, and Danny M. Jasper. 2014. "Changing Race Boundary Perception by Reading Narrative Fiction," *Basic and Applied Social Psychology* 36(1): 83–90. https://doi.org/10.1080/01973533.2013.856791

Johnson, Mark. 1993. *Moral Imagination: Implications of Cognitive Science for Ethics*. Chicago: University of Chicago Press.

Johnson, Mark. 2014. *Morality for Humans: Ethical Understanding from the Perspective of Cognitive Science.* Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226113548.001.0001

Kang, Yoona, John Dovidio, and Jeremy Gray. 2013. "The Nondiscriminating Heart: Lovingkindness Meditation Training Decreases Implicit Intergroup Bias," *Journal of Experimental Psychology*: 1–8.

Kenyon, Tim. 2014. "False Polarization: Debiasing as Applied Social Epistemology," *Synthese* 191(11): 2529–47. https://doi.org/10.1007/s11229-014-0438-x

Kenyon, Tim, and Guillaume Beaulac. 2014. "Critical Thinking Education and Debiasing," *Informal Logic* 34(4): 341–63. https://doi.org/10.22329/il.v34i4.4203

Kretz, Lisa. 2012. "Climate Change: Bridging the Theory-Action Gap," *Ethics & the Environment* 17(2): 9–27. https://doi.org/10.2979/ethicsenviro.17.2.9

Kretz, Lisa. 2014. "Teaching Being Ethical," *Teaching Ethics* 15(1): 151–72. https://doi.org/10.5840/tej201515112

Lai, Calvin K., Maddalena Marini, Steven A. Lehr, Carlo Cerruti, Jiyun-Elizabeth L. Shin, Jennifer A. Joy-Gaba, Arnold K. Ho, Bethany A. Teachman, Sean Wojcik, Spassena P. Koleva, Rebecca S,. Frazier, Larisa Heiphetz, Eva E. Chen, Rhiannon N. Turner, Jonathan Haidt, Selin Kesebir, Carlee Beth Hawkins, Hillary S. Schaefer, Sandro Rubichi, Giuseppe Sartori, Christopher M. Dial, N. Sriram, Mahzarin R. Banaji, and Brian A. Nosek. 2014. "Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions," *Journal of Experimental Psychology: General* 143(4): 1765–85. https://doi.org/10.1037/a0036260

Lai, Calvin K., Allison L. Skinner, Erin Cooley, Sohad Murrar, Markus Brauer, Thierry Devos, Jimmy Calanchini, Y. Jenny Xiao, Christina Pedram, Christopher K. Marshburn, Stefanie Simon, John C. Blanchar, Jennifer A. Joy-Gaba, John Conway, Liz Redford,

Rick A. Klein, Gina Roussos, Fabian M. H. Schellhaas, Mason Burns, Xiaoqing Hu, Meghan C. McLean, Jordan R. Axt, Shaki Asgari, Kathleen Schmidt, Rachel Rubinstein, Maddalena Marini, Sandro Rubichi, Jiyun-Elizabeth L. Shin, and Brian A. Nosek. 2016. "Reducing Implicit Racial Preferences: II. Intervention Effectiveness across Time," *Journal of Experimental Psychology* 145(8): 1001–16. https://doi.org/10.1037/xge0000179

Lorde, Audre. 1984. "Age, Race, Class, and Sex: Women Redefining Difference," in *Sister Outsider: Essays and Speeches*, 114–23. Berkeley, Calif.: The Crossing Press.

Lueke, Adam, and Bryan Gibson. 2014. "Mindfulness Meditation Reduces Implicit Age and Race Bias: The Role of Reduced Automaticity of Responding," *Social Psychology and Personality Science*: 1–8.

Lueke, Adam, and Bryan Gibson. 2016. "Brief Mindfulness Meditation Reduces Discrimination," *Psychology of Consciousness: Theory, Research, and Practice* 3(1): 34–44. https://doi.org/10.1037/cns0000081

Mar, R. A. 2011. "The Neural Bases of Social Cognition and Story Comprehension," *Annual Review of Psychology* 62: 103–34. https://doi.org/10.1146/ annurev-psych-120709-145406

Maxwell, Bruce. 2016. "The Debiasing Agenda in Ethics Teaching: An Overview and Appraisal of the Behavioral Ethics Perspective," *Teaching Ethics* 16(1): 75–90. https://doi.org/10.5840/tej20164132

Morewedge, Carey K., Haewon Yoon, Irene Scopelliti, Carl W. Symborski, James H. Korris, and Karim S. Kassam. 2015. "Debiasing Decisions: Improved Decision Making with a Single Training Intervention," *Policy Insights from the Behavioral and Brain Sciences* 2(1): 129–40. https://doi.org/10.1177/2372732215600886

Mower, Deborah. 2008. "Teaching Sympathetic Moral Reasoning," *Teaching Ethics* 8(2): 1–14. https://doi.org/10.5840/tej2008821

Mower, Deborah. 2015. "Reflections on . . . A Culture of Sensitivity," *Teaching Ethics* 15(1): 1–18. https://doi.org/10.5840/tej20151512

Nisbett, Richard E., and Timothy D. Wilson. 1977. "Telling More than We Can Know: Verbal Reports on Mental Processes," *Psychological Review* 84(3): 231–59. https://doi.org/10.1037/0033-295X.84.3.231

Paluck, Elizabeth L., and Donald P. Green. 2009. "Prejudice Reduction: What Works? A Review and Assessment of Research and Practice," *Annual Review of Psychology* 60: 339–67. https://doi.org/10.1146/annurev.psych.60.110707.163607

Project Implicit. https://implicit.harvard.edu/implicit/. Accessed May 2016.

Railton, Peter. 2014. "The Affective Dog and its Rational Tale: Intuition and Attunement," *Ethics* 124(4): 813–59. https://doi.org/10.1086/675876

Rees, Clea F.. 2016. "A Virtue Ethics Response to Implicit Bias," in *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, ed. Michael Brownstein and Jennifer Saul, 191–214. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198766179.003.0009

Ricard, Matthieu. 2015. *Altruism*. New York: Little, Brown and Company.

Rudman, L. A., R. D. Ashmore, and M. L. Gary. 2001. "'Unlearning' Automatic Biases: The Malleability of Implicit Prejudice and Stereotypes," *Journal of Personality and Social Psychology* 81: 856–68. https://doi.org/10.1037/0022-3514.81.5.856

Sherwin, Susan. 1999. "Foundations, Frameworks, Lenses: The Role of Theories in Bioethics," *Bioethics* 13(3–4): 198–205. https://doi.org/10.1111/1467-8519.00147

Singer, Peter. 2002. *Animal Liberation*. New York: Harper Collins Publishers Inc.

Speer, N. K., J. R. Reynolds, K. M. Swallow, and J. M. Zacks. 2009. "Reading Stories Activates Neural Representations of Visual and Motor Experiences," *Psychological Science* 20: 989–99. https://doi.org/10.1111/j.1467-9280.2009.02397.x

Stell, Alexander J., and Tom Farsides. 2016. "Brief Loving-Kindness Meditation Reduces Racial Bias, Mediated by Positive Other-Regarding Emotions," *Motivation and Emotion* 40(1): 140–47. https://doi.org/10.1007/s11031-015-9514-x

Stephan, W. G., and K. Finlay. 1999. "The Role of Empathy in Improving Intergroup Relations," *Journal of Social Issues* 554: 729–44.

Stout, J. G., N. Dasgupta, M. Hunsinger, and M. A. McManus. 2011. "STEMing the Tide: Using Ingroup Experts to Inoculate Women's Self-concept in Science, Technology, Engineering, and Mathematics (STEM)," *Journal of Personality and Social Psychology* 100: 255–70. https://doi.org/10.1037/a0021385

Thomas, Laurence. 1992–1993. "Moral Deference," *The Philosophical Forum* 14: 233–50.

Tuana, Nancy. 2007. "Conceptualizing Moral Literacy," *Journal of Educational Administration* 45(4): 364–78. https://doi.org/10.1108/09578230710762409

Urban Walker, Margaret. 1997. *Moral Understandings: A Feminist Study in Ethics*. London: Routledge.

Warren, Karen. 2005. "The Power and Promise of Ecofeminism, Revisited," in *Environmental Philosophy: From Animal Rights to Radical Ecology*, ed. Michael Zimmerman, J. Baird Callicott, Karen Warren, Irene Klaver, and John Clark, 252–79. Upper Saddle River, N.J.: Prentice Hall.

Washington, Natalia, and Daniel Kelly. 2016. "Who's Responsible for This? Moral Responsibility, Externalism and Knowledge about Implicit Bias," in *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, ed. Michael Brownstein and Jennifer Saul, 11–36. Oxford: Oxford University Press.

Wilke, A., and R. Mata. 2012. "Cognitive Bias," in *Encyclopedia of Human Behavior*, ed. V. S. Ramachandran. London: Academic Press.

Wilson, Timothy D., and Nancy Brekke. 1994. "Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations," *Psychological Bulletin* 116(1): 117–42. https://doi.org/10.1037/0033-2909.116.1.117

Yap, Audrey. 2013. "Ad Hominem Fallacies, Bias and Testimony," *Argumentation* 27(2): 97–109. https://doi.org/10.1007/s10503-011-9260-5

Zajonc, R. B. 1980. "Feeling and Thinking: Preferences Need No Inferences," *American Psychologist* 35(2): 151–75. https://doi.org/10.1037/0003-066X.35.2.151

Lisa Kretz is an assistant professor of philosophy at the University of Evansville. Her recent and forthcoming publications are at the intersection of areas such as ecological selfhood, the theory-action gap, motivational framing for action, ecological emotions (with a special emphasis on hope), moral psychology, the politics of emotion, and activist pedagogy. She also has a number of publications that identify and argue against the immoral treatment of non-human animals. Lisa Kretz, Department of Philosophy, University of Evansville, Evansville IN 47714; lk102@evansville.edu