

INSTRUCTOR'S EDITION

This free instructor's copy has been provided only for use in textbook evaluation. Sale or resale is prohibited. This book remains the property of Oxford University Press.

COLLEGE ETHICS

A READER ON MORAL ISSUES
THAT AFFECT YOU

Compare the price and quality of our book to
other books published for your course:

Publisher	Title	Author	Price
Oxford November 2016	College Ethics	Fischer	\$69.95
McGraw-Hill May 2011	Moral Ethics: Morality and Social Policy, 3e	Mappin	\$202.67
Chicago January 2011	Contemporary Moral Problems, 10e	Wright	\$198.95
Wiley-Blackwell September 2012	Morally and Moral Controversies, 9e	Scott/Walton	\$115.96

*List (suggested retail) prices as advertised on publishers' websites as of September 2016. Prices of cited are for comparable formats. Other formats may be available. Prices subject to change.

Oxford University Press USA is a not-for-profit publisher. For more information, please call 800-280-0280.

Bob Fischer

OXFORD
UNIVERSITY PRESS

Racial Cognition and the Ethics of Implicit Bias

Daniel Kelly and Erica Roedder

.....
Daniel Kelly and Erica Roedder examine the nature and morality of implicit racial bias in their essay. First they review some of the evidence for the idea that implicit racial bias, that is, the implicit association of negative traits with people of certain races, is a widespread phenomenon and one that may influence a person's behavior. Next, Kelly and Roedder delve into the problem of how to morally assess implicit racial bias. They suggest that there are good reasons for thinking that even having such a bias is morally wrong, and that if you can reasonably suppose that you're biased, you should attempt to correct for it in practice, even in the absence of empirical evidence for it.
.....

1. Introduction

...
The aim of this paper is two-fold. Our first goal is to call philosophical attention to some of the most provocative empirical work on racial cognition. Accordingly, the first half of the paper will discuss one portion

Kelly, Daniel, and Erica Roedder. 2008. "Racial Cognition and the Ethics of Implicit Bias." *Philosophy Compass* 3(3): 522–540. Reprinted with permission of John Wiley and Sons.

of this large literature: work regarding implicit racial biases. Our second goal is to raise a number of philosophical questions about the proper normative assessment of behaviors and judgments linked to those implicit biases. In the second half of the paper, then, we will assume these implicit racial biases are roughly as current research depicts them to be, and go on to sketch a few of the most promising avenues of philosophical research that we believe are opened up by the psychological complexities revealed in this work on racial cognition.

2. Implicit Racial Bias

Rather ingenious strategies have uncovered subtle forms of racial discrimination that still exist in real world settings. One recent study investigated the effect of race on hiring practices in two U.S. cities. Researchers sent out fabricated résumés to Help Wanted ads appearing in major newspapers in Boston and Chicago. Half of the résumés were headed by a very Black sounding name (e.g., Lakisha and Jamal), while the other half were headed by a very White sounding name (e.g., Emily and Greg).¹ The results were remarkable: overall, résumés bearing White names received an astonishing 50 percent more callbacks for interviews than their Black counterparts. Furthermore, an interesting pattern emerged for highly qualified résumés. For White sounding names, résumés with highly qualified credentials received 30 percent more callbacks than their less qualified White counterparts; in contrast, employers did not differentiate nearly as much between highly qualified Black résumés and their less qualified Black counterparts. The amount of discrimination was fairly consistent across occupations and industries. Of particular interest was the fact that employers who explicitly listed "Equal Opportunity Employer" in their ad discriminated just as much as other employers (Bertrand and Mullainathan).

... One of the most sophisticated and widely used windows into racial cognition is an experimental measurement technique called the Implicit Association Test, or IAT for short. More than any other technique, the IAT has been used to establish the existence and shed light on the character of implicit racial biases. In short, the IAT has been used to show that a great many people, including those who genuinely profess themselves to be racially impartial and explicitly disavow any form of racial prejudice, display subtle signs of racial bias in controlled experimental settings. ...

The Implicit Association Test (IAT)

The IAT was designed by psychologists to probe aspects of thought that are not easily accessible or immediately available to introspection.² ...

... The core idea behind ... IATs is that stronger associations between items will allow them to be grouped together more quickly and easily.³ For instance, faster and more accurate performance on iterations when good and White items are to be grouped together than on iterations when good and Black are to be grouped together indicates a stronger association between good and White. Stronger associations between good and White, in turn, are taken to indicate an implicit bias towards Whites over Blacks. As should be evident, this test does not use self-report or explicitly ask subjects about their attitudes about race. Unlike those more direct tests that are based on self report, and which are often used in conjunction with IATs (e.g. McConahay), the IAT requires subjects to make snap judgments that must be made quickly, and thus without moderating influence of introspection and deliberation and often without conscious intention. Biases revealed by an IAT are often thought to implicate relatively automatic processes.

IAT and Race

... In using tools like the IAT in conjunction with more direct, self-report methods, researchers have further found that even those who sincerely profess tolerant or anti-racist views can nevertheless harbor implicit racial biases (often to their own surprise and chagrin).⁴ Counterintuitive as it may seem, this robust pattern of results shows that a person's avowed views on race and racism are not a reliable guide to whether or not they are implicitly biased.

... We will conclude with a final example that speaks to both the influence of IAT results on behavior and real world relevance. ... many important judgments must be made almost instantaneously and in high pressure situations. Such split second decisions have been shown to be sensitive to race in other ways as well. A number of studies have asked people to make snap decisions about whether a presented object is a gun or some other harmless object. Researchers found that when first shown a picture of a Black face, both White and Black Americans become more likely to misidentify a harmless object as a gun (Payne, "Weapon Bias"). Not only is this "weapon bias" found in people who explicitly try to avoid racial biases, but the weapon bias is highly correlated with the indirect measures of racial biases, including the IAT (Payne, "Conceptualizing

Control"). The relevance of such findings is difficult to deny, especially in light of tragedies such as the 1999 shooting of Amadou Diallo, who was shot forty-one times by police officers who thought he was drawing a gun; he was actually just reaching for his wallet.

3. Normative Questions

So far, we have discussed the psychology of racial cognition, focusing on the implicit attitude test. Such findings introduce new and significant normative questions. In the rest of this article, we'll briefly survey some of the normative questions that we think are fruitful areas for future research on racial cognition, and consider attempts to answer questions similar to them. . . .

Is It Morally Problematic to Harbor Implicit Racial Bias?

One major question is whether it is morally problematic, in and of itself, to have an implicit bias against members of a particular race.⁵ Obviously, implicit racial bias is problematic insofar as it leads to harmful or unfair consequences. For instance, suppose implicit bias forms part of the explanation of why an innocent Black man is shot by a police officer. In this case, implicit bias is clearly a bad thing: It partly caused a *harmful* consequence, i.e., the death of a young man. Similarly, implicit racial bias is clearly bad insofar as it leads to *unfair* consequences, e.g., the unequal promotion of White versus Black employees within a company.

Let us set aside such consequences for a moment and consider the question of the implicit attitude itself—is this attitude intrinsically a bad thing? . . .

. . . consider an *explicitly* racist person. We might ask of him, is his explicitly racist attitude, in and of itself, a bad thing? Suppose, for instance, that a man were never to act on his explicit racial beliefs, keeping his racist thoughts and feelings to himself. Perhaps he secretly seethes with disgust after drinking from water fountains used by Blacks and often has thoughts like, "It's so obvious that Black children aren't as smart as White children."⁶

Most Westerners, we suspect, would disapprove of such a person. Even if the man never acts on these racist thoughts and feelings, and even if he is morally upright in all the other aspects of his life (e.g., he goes to church, is faithful to his wife, etc.), there is still something morally problematic about his attitudes. While it's good that the man refrains from

acting on these racist thoughts and feelings, it is unfortunate and morally condemnable that he has such attitudes at all.

. . . If this is right, can the same be said about *implicit* thoughts and feelings? For instance, what exactly *are* implicit attitudes? Are they akin to Freudian unconscious states, occupying some deep core of our psyche? Or are they more minimal and peripheral? . . . Let us consider, for a moment, an extremely minimal construal of implicit attitudes suggested by this: an implicit attitude is simply a tendency to associate one concept with another, in the way that, for instance, the concept *salt* might prime the concept *pepper*. A high IAT score, on this understanding, means that a person strongly associates, e.g., Black faces with handguns. Assuming that this is an exhaustive description of the implicit attitude—a tendency to associate one concept with another—can a tendency to associate certain concepts, in and of itself, be morally problematic?

One way to approach this question is through the lens of rationality. While it is clear that explicitly racist beliefs are mostly irrational, in addition to being immoral (e.g., the thought that Black children are less smart than White children), there seems to be room to argue that some implicit racial associations are (to a *limited extent*) rational. . . .

To see why someone might argue that implicit attitudes are sometimes rational, let's first consider a different case, i.e., gender. IAT results suggest that most people strongly associate men with science, more so than they do women with science (see Nosek and Banaji). But if the implicit attitude really is *just* an association of concepts, might it be rational to make such associations? Women, as a matter of fact, are not as well-represented in the sciences. . . . With respect to the issue of rationality, our point is that if implicit attitudes are construed in this very minimal way—as indicating only that a person associates two concepts—it appears they can be rational in some sense (e.g., insofar the association between concepts accurately reflects a correlation or statistical regularity that holds among those referents of the concepts).

Let us now return to the racial example. Consider the tendency to associate the faces of young, Black men with handguns. Someone might analogously suggest that, were it true that young Black men carry guns at a higher rate than White men, then it would be rational to associate Black faces with handguns. This is important because, as we mentioned earlier, it might be thought that rationality and morality go hand-in-hand: . . .

We suspect this is not the right way to think about rationality and implicit attitudes. First, we think that a rational attitude may still be an *immoral* one. Rationality and morality are different virtues, so it should be expected that a person can have the one without the other. . . .

Our point is that it can sometimes be unkind or uncompassionate to believe ill of a person, even if it is rational to do so. Thus it can sometimes be immoral to hold a belief that is, in fact, rational.

As a second point, suppose we were to grant, for the sake of argument, the suggestion that rational attitudes are moral and irrational ones are immoral. Even on this supposition, a case can still be made that implicit racial biases are morally problematic. We suspect that such associations (such as those found in studies on the weapon biases) almost always extend beyond what is rational, and there will almost always be a "remainder": an implicit association that goes beyond what rationality endorses. If this is right, then even on the supposition that morality and rationality are tightly bound together, implicit attitudes will remain morally problematic to the extent that they outstrip what is rationally justifiable. . . .

. . . In the remainder of this section, we'll examine . . . the work of two authors: Garcia's account of racism and Blum's account of stereotyping.

Garcia's analysis of racism stresses the intrinsic features of certain attitudes. He writes that someone is a racist when they have certain affective and volitional attitudes:

[W]hat makes someone a racist is her disregard for, or even hostility to, those assigned to the targeted race . . . [S]he is hostile to or cares nothing (or too little) about some people because of their racial classification . . . [H]ate and callous indifference (like love) are principally matters of *will* and desire: what does one want, what would one choose, for those assigned to this or that race? (43)

Importantly, Garcia construes racism as a deformation of affect and the will, and this informs his account of why it is morally problematic: racist attitudes, in themselves, are "inherently contrary to the moral virtues of benevolence and justice" (43). Such attitudes, he argues, are hateful and ill-willed, and are thus opposed to benevolence by their very nature. On Garcia's account, the question of whether it is wrong to harbor an *implicit* attitude will therefore boil down to whether the attitude is intrinsically opposed to benevolence, e.g., whether it is an attitude of hate or one of ill will.

Determining whether implicit attitudes are intrinsically opposed to benevolence, however, will require progress on two fronts. First, there are

issues tied to empirical work and how to interpret evidence provided by indirect tests. Implicit attitudes (or some implicit attitudes) may turn out to be *merely cognitive* associations, in which case they would be neither affective nor volitional. Such attitudes, on Garcia's account, would not be intrinsically opposed to benevolence, and so would not be morally problematic.⁷

Suppose, on the other hand, some implicit attitudes are indeed affectively laden, as a growing body of empirical research suggests (e.g., Vanman et al.; Phelps et al.; Amodio, Harmon-Jones, and Devine; see also Payne et al.). This possibility raises a different kind of difficulty, which turns on whether such implicit attitudes should be thought of as "inherently contrary to the moral virtue of benevolence." While it is obvious that explicit, hate-filled racial rage is intrinsically opposed to benevolence, it is far less clear whether the more subtle attitudes measured by the IAT ought to be categorized in this way. . . .

Let us turn now to Blum's account of racial stereotyping. . . .

In much of his article, Blum analyzes what stereotypes *do*. Two of the most important features he describes are that they mask individuality (the stereotyper fails to be sensitive to an individual's quirks and characteristics) and that they lead to what he calls *moral distancing*. In *moral distancing*, the stereotyper sees a stereotypee as more "other" than he or she really is; and this corrodes her sense of a common, shared humanity. Here, we think Blum's account can be usefully and straightforwardly generalized to implicit attitudes. One must simply ask: do implicit attitudes have these deleterious effects? Do implicit biases mask individuality and lead to moral distancing? These sound like clear-cut empirical questions. If implicit racial biases do lead thinkers to fail to appreciate the individuality of others or to morally distance themselves, then it follows from Blum's account that those implicit biases are morally reprehensible.

As a final note, it seems to us that ethicists working on implicit racism might be well-served by making a distinction between what is wrong and what is morally blameworthy. Particularly in the case of implicit attitudes, it is salient that their acquisition may be rapid, automatic, and uncontrollable.⁸ These features, it might be thought, are related to features that establish blameworthiness—such as identification (Frankfurt) or reasons-responsiveness (Fischer and Ravizza). For instance, it might be said that the implicitly racist person doesn't identify with his implicit attitude, or that the attitude isn't responsive to reasons; thus we cannot hold a person fully accountable for those implicit attitudes. If this is right, one

might say that such attitudes are morally wrong—and condemnable—but that the person himself cannot be blamed for having them. . . .

Might I Be Racially Biased?

One of the remarkable features of *implicit* bias is the possibility that individuals may not be aware of their own bias. Neither introspection nor honest self-report are reliable guides to the presence of such mental states, and one may harbor implicit biases that are diametrically opposed to one's explicitly stated and consciously avowed attitudes. Because of this, thinkers face a thorny, real-life epistemological problem: Given that a large proportion of the population is implicitly racially biased, is it reasonable to conclude that I, myself, am racially biased? And if I believe I might be, how should that belief affect my deliberation and behavior?

The possibility that you, yourself, may harbor implicit biases has implications for your concrete beliefs about everyday matters. For instance, suppose you are a White professor grading a Black student's paper, and you are initially inclined to give the paper an 89/100. Does the possibility of implicit racial bias give you good reason to think the paper actually deserves slightly better, e.g., 90 or 91 points? Let's call this example *the savvy grader*, since the problem arises when a thinker is psychologically savvy and is thus aware of the prevalence of implicit racial bias (the example is discussed in Roedder).

An analogy will be helpful here. Suppose you learn of psychological research showing that most people are inclined to underestimate the size of circles when set across a hatched background. Suppose you are later asked to judge the size of a circle on a hatched background. In deciding the size of the circle, it is most rational to estimate it to be slightly larger than you are initially inclined to guess. . . .

With this in mind, let us return to the case of the savvy grader. Assume for a moment that experiments uncovered a racial bias in the grading of student papers. . . . We maintain that by parity of reasoning, it would be wise to make a similar adjustment for the implicit bias in grading, just as you would correct for the visual bias in judging the size of a circle.⁹ In both cases, one is acting for purely *epistemic* reasons; in order to give the most accurate grade, i.e., in order to grade the paper based on its merits, it is reasonable for the savvy grader to correct for the effects of racial biases.¹⁰

There is much more to say here. In particular, we might wonder how much evidence of implicit racial bias a savvy grader needs before it is reasonable for her to adjust how she assigns grades. Roedder argues that the epistemic requirements are strikingly low: It is enough if she knows that, *ceteris paribus*, the bias exists *on average*. Consider the visual analogy again. If one were told that, *on average*, people see the circle as 25 percent smaller than it really is, most of us would take that as a reason to increase our original estimate of its size by 25 percent. Here, too, the epistemic factors relevant to grading papers do not appear substantially different from those of the visual case. . . .

Of course, we don't yet have evidence that directly bears on the question of whether or not normal thinkers are implicitly biased against their Black students when grading papers; to date there has not been a systematic effort to look for racial bias in essay grading at the college level. . . .

Thus the important question is this: Knowing what you know now about implicit bias in other domains (perhaps from reading this very article!), and if you had to place a bet, would you bet that there *is* a racial bias in grading or that there *isn't*?¹¹ If you find yourself inclined to think that (more likely than not) there is a racial bias in grading, and if the line of reasoning sketched here is correct, then merely having this empirical hunch is enough to rationally compel you to make some sort of compensatory adjustment in your Black students' grades.¹² . . .

. . . Implicit racial bias is just one example where psychological science shows our *reasoning* capacities to be impaired, and where we have *no introspective access to our own impairment*. Whenever this is the case, and wherever thinkers are savvy enough to learn about the psychology of such biases, similar epistemological challenges concerning self-assessment and proper adjustment are likely to arise.

References

- Amadio, D., E. Harmon-Jones, and P. Devine. (2003). "Individual Differences in the Activation and Control of Affective Race Bias as Assessed by Startle Eyeblink Response and Self-Report." *Journal of Personality and Social Psychology* 84 (4), 738–53.
- Blum, Lawrence. (2004). "Stereotypes and Stereotyping: A Moral Analysis." *Philosophical Papers* 3, 251–89.
- Cunningham, W., K. Preacher, and M. Banaji. (2001). "Implicit Attitude Measures: Consistency, Stability, and Convergent Validity." *Psychological Science* 12 (2), 163–70.

- Devine, P., et al. (2002). "The Regulation of Explicit and Implicit Race Bias: The Role of Motivations to Respond Without Prejudice." *Journal of Personality and Social Psychology* 82 (5), 835–48.
- Fischer, J. and M. Ravizza. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Garcia, J. L. A. (2004). "Three Sites for Racism: Social Structures, Valuations and Vice." In *Racism in Mind*, eds. M. P. Levine and T. Pataki. Ithaca, NY: Cornell University Press, 36–55.
- Greenwald, A., D. McGhee, and J. Schwartz. (1998). "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74 (6), 1464–80.
- Greenwald, A. and A. Nosek. (2001). "Health of the Implicit Association Test at age 3." *Zeitschrift für Experimentelle Psychologie* 48, 85–93.
- McCónahay, J. (1986). "Modern Racism, Ambivalence, and the Modern Racism Scale." In *Prejudice, Discrimination, and Racism*, eds. J. F. Dovidio and S. L. Gaertner. Orlando, FL: Academic Press.
- Nosek, B. A., A. G. Greenwald, and M. R. Banaji. (2007). "The Implicit Association Test at Age 7: A Methodological and Conceptual Review." In *Automatic Processes in Social Thinking and Behavior*, ed. J. A. Bargh. Philadelphia, PA: Psychology Press.
- Nosek, B. A., A. G. Greenwald, and M. R. Banaji. (2005). "Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity." *Personality and Social Psychology Bulletin* 31 (2), 166–80.
- Payne, B. K. (2005). "Conceptualizing Control in Social Cognition: The Role of Automatic and Controlled Processes in Misperceiving a Weapon." *Journal of Personality Social Psychology* 81, 181–92.
- Payne, B. K. (2006). "Weapon Bias: Split-Second Decisions and Unintended Stereotyping." *Current Directions in Psychological Science* 15, 287–91.
- Payne, B. K., et al. (2005). "An Inkblot for Attitudes: Affect Misattribution as Implicit Measurement." *Journal of Personality and Social Psychology* 89 (3), 277–293.
- Phelps, E., et al. (2000). "Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation." *Journal of Cognitive Neuroscience* 12 (5), 729–38.
- Roedder, E. (2010). "Savvy Thinking" and "The Epistemology of Self-Correction for Implicit Bias." *Beings Like Us: Deliberating in Light of Psychological Theory*. Ph.D. dissertation, New York University, Department of Philosophy.
- Vanman, E. J., et al. (1997). "The Modern Face of Prejudice and Structural Features that Moderate the Effect of Cooperation on Affect." *Journal of Personality and Social Psychology* 73, 941–59.

Comprehension Questions

1. What does the study about the fabricated résumés show about people's attitudes regarding race?

2. How does the Implicit Association Test work, and what does it tell us about people's attitudes regarding race?
3. Why might it be wrong to have an explicit racist attitude, even if you don't act on it?
4. In what sense might some implicit racial biases be rational?
5. Regardless of its rationality, why do you think that it's wrong to have implicit racial biases?
6. What reason do Kelly and Roedder give for thinking that you should correct for implicit racial bias in yourself, given adequate evidence of it?
7. Why do Kelly and Roedder think you ought to correct for implicit racial bias in yourself even without having adequate evidence for it?

Discussion Questions

1. Kelly and Roedder focus on implicit racial bias and discuss the case of a white instructor making corrections for that bias when evaluating the work of black students. What types of implicit bias might you be susceptible to, and do you think you ought to correct for it somehow?
2. There are several different reasons discussed by Kelly and Roedder for thinking having implicit racial biases is morally wrong. Which do you think is the strongest? Can you think of any additional reasons?
3. Kelly and Roedder suggest that even if it is morally wrong to have implicit racial biases, it might nevertheless not be morally blameworthy to have them. Do you think the distinction between wrongness and blameworthiness should apply in this case?

Case

Here's a summary of a troubling study:

Two hundred ten Princeton University undergraduates . . . were assigned to one of 105 same-sex pairs. A white student was assigned the role of instructor; a second student, who could be either black or white, played the student. "Instructors first completed an ostensible cognitive flexibility task that, in reality, was a subliminal priming task inconspicuously assessing implicit racial bias," the researchers write. . . . The teachers were supplied with materials for creating their lesson. . . . The teachers were supplied with materials for creating their lesson—on Byzantine history—and given eighteen minutes to prepare. The instructors then met their students, and proceeded to teach a seven-minute lesson. Two coders watched videotapes of these encounters

and graded each teacher for nervousness, discomfort, awkwardness, and stiffness. They also rated the quality of the lesson, grading it for, among other things, clarity and the coherence of the material.

Following the lesson, while the instructor completed a test designed to measure conscious, explicit bias, the student took a test on the material he or she was just taught, which included recalling important names, dates, and other relevant facts.

The result: Once explicit bias was removed from the equation, greater unconscious bias on the part of an instructor predicted lower test scores for black, but not white, students. "Our evidence suggests this was due to higher bias instructors being more anxious, and therefore giving less effective lessons," the researchers write.¹³

What should we say about these instructors in light of Kelly and Roedder's arguments? What should instructors do with this information? What should students do with it?

Notes

1. Throughout, we will simplify the discussion by considering just two groups, and using the capitalized terms "Black" and "White" to refer to those putative racial groups and their members. Other terminology, e.g., "African-American," is less suitable for our purposes because it is overly restrictive. For example, it does not appear that implicit racial biases against Blacks apply only to Black Americans, or only to Americans of specifically African descent.
2. See Greenwald, McGhee, and Schwartz for the first presentation of the IAT itself, as well as the initial results obtained with it. Also see Greenwald and Nosek; Lane et al.; Nosek, Greenwald, and Banaji, "Implicit Association Test" for more recent reviews of data gathered using IATs, and for useful discussions of the methodological issues surrounding the test.
3. More precisely: "the logic of the IAT is that this sorting task should be easier when the two concepts that share a response are strongly associated than when they are weakly associated" (Nosek, Greenwald, and Banaji, "Implicit Association Test" 267).
4. Similar dissociations have been found using a wide variety of other indirect measures, including evaluative priming (Cunningham, Preacher, and Banaji; Devine et al.), the startle eyeblink test (Phelps et al.; Amodio, Harmon-Jones, and Devine), and EMG measures (Vanman et al.).
5. We know of no efforts to answer this question, although it is posed in Jolls and Sunstein.
6. In stating these examples, we felt *extremely* uncomfortable, and we anticipate that our readers will feel the same way. However, we think concrete examples are needed in order to make salient our point: In general, explicitly racist attitudes—even if they are not acted upon—are morally damnable.
7. At least, they would not be morally problematic *in the way* that racist attitudes are problematic. Garcia offers an account of racism, not a complete moral theory.
8. See Gregg, Seibt, and Banaji. We have stated that it is more *salient* that implicit attitudes are uncontrollable. That's because, arguably, the acquisition of most *explicit* attitudes is uncontrollable as well; it's just not salient at first glance. One does not control one's acquisition of, for instance, one's beliefs about plants, one's attitudes towards pets, etc. So one will need to appeal to more complex or carefully delineated features—perhaps identification or reasons-responsiveness—if one wants to claim that implicit attitudes are not proper subjects of blame, but that explicit attitudes are.
9. There are, of course, many ways one might go about compensating for implicit racial bias. Most obviously, one might use conscious rules, e.g., "Bump up borderline grades of Black students." In addition, there are various psychological techniques which seem to mitigate implicit racial bias, such as entertaining counterstereotypic thoughts (e.g., imagining a positive Black role model, or a female scientist).
10. People sometimes question the idea that grades are apt to be "accurate" at all. In many ways, this is irrelevant—most of us want to avoid having the race of a student affect their grading. That desire is enough to motivate the problems we raise here: If one has this desire, it seems that one is rationally compelled to correct for possible influence of race in deliberation.
11. To add to the case we are making, one can appeal to expert opinion. In that vein, we have discussed this issue with two members of Banaji's lab at Harvard, both of whom said they'd be "very surprised" if there wasn't implicit racial bias in the domain of grading.
12. This version of epistemic argument is highly compressed, thus there isn't room to respond to a number of important objections. Roedder contains a much fuller exploration and defense of these claims.
13. <http://www.psmag.com/books-and-culture/unconscious-teacher-bias-harms-black-college-students>