A lesson in bias: The relationship between implicit racial bias and performance in pedagogical contexts

Drew S. Jacoby-Senghor, Stacey Sinclair, J. Nicole Shelton

PII: S0022-1031(15)30010-X
DOI: doi: 10.1016/j.jesp.2015.10.010
Reference: YJESP 3378

To appear in: Journal of Experimental Social Psychology

Received date: 23 June 2015
Revised date: 27 October 2015
Accepted date: 28 October 2015

Please cite this article as: Jacoby-Senghor, D.S., Sinclair, S. & Nicole Shelton, J., A lesson in bias: The relationship between implicit racial bias and performance in pedagogical contexts, Journal of Experimental Social Psychology (2015), doi: 10.1016/j.jesp.2015.10.010

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
A Lesson in Bias:

The Relationship Between Implicit Racial Bias and Performance in Pedagogical Contexts

Drew S. Jacoby-Senghor¹

Stacey Sinclair²

J. Nicole Shelton²

¹Columbia University

²Princeton University

Word count: 2969

Corresponding Author: Drew S. Jacoby-Senghor, Uris Hall, 3022 Broadway, New York, NY 10027-6902 E-mail: dsenghor@gmail.com
Abstract

We posit instructors’ implicit racial bias as a factor in racial disparities in academic achievement and test the relationship between this factor, instructor lesson quality, and learners’ subsequent test performance. In Study 1, white participants were assigned to the role of instructor and gave a short lesson to a learner who was either black or white. Instructors’ implicit bias predicted diminished test performance on the part of black, but not white, learners. Further, instructors’ anxiety and lesson quality, as rated by coders, mediated the relationship between their implicit bias and learners’ test performance. In Study 2, a separate sample of non-black participants watched videos of instructors from cross-race lessons from the first experiment. Once again, instructors’ implicit bias predicted diminished test performance by participants. These findings suggest that underperformance by minorities in academic domains may be driven by the effect implicit racial biases have on educators’ pedagogical effectiveness.

Keywords: implicit prejudice, intergroup interaction, educational disparities, attitudes
A Lesson in Bias: The Relationship Between Implicit Racial Bias and Performance in Pedagogical Contexts

Over seventy years after the U.S. Supreme Court declared segregation in public schools unconstitutional, widespread disparities persist between the educational outcomes of blacks and whites (Nord et al., 2011). Modern social psychology has increasingly focused on minority students’ responses to being stigmatized as a causal factor of these disparities. Substantial research on identity threats shows that the perception of bias can preoccupy minorities, siphoning away cognitive resources needed to optimally perform (Steele, Spencer, & Aronson, 2002). For example, Logel and colleagues (2009) show that female engineering students who interact with subtly sexist, versus nonsexist, men experience identity threat that causes worse performance on engineering tests. Aside from concerns about being stereotyped, minorities’ intellectual engagement and performance are hindered by the mere concern that they do not belong (Walton & Cohen, 2007) and may be socially rejected (Mendoza-Denton, Downey, Purdie, Davis, & Pietrzak 2002) because of their group membership. Together, this literature shows that the possibility or perception of prejudice can loom large in the minds of minorities, hampering their academic achievement.

While research from this perspective has increased dramatically, corresponding examination of how teachers contribute to the problem has waned. Expectancy confirmation research remains the most prominent perspective (e.g., Word, Zanna, & Cooper, 1974), showing that teachers’ biased expectations for their minority students negatively impact students’ outcomes (Rubovits & Maehr, 1973; Weinstein, Gregory, & Strambler, 2004). However, the role that expectancy confirmation plays in modern academic disparities has been criticized on several dimensions, including arguments that expectancies relate to outcomes because they are actually
accurate (Jussim, 1989; Jussim & Eccles, 1995; Jussim, Eccles, Madon, 1996) and that they do not bias instructors’ evaluative judgments (Trouillard, Sarrazin, Martinek, Guillet, 2002). Furthermore, historic reductions in explicit racial bias (Bobo & Charles, 2009) make it less likely that teachers undermine students’ progress because of intentionally prejudiced expectancies.

In contrast, the present research investigates whether teachers’ unintentional and automatic (i.e. implicit) racial bias produces racial disparities in academic achievement. Among whites, greater implicit prejudice is associated with heightened anxiety during interracial interaction; it is positively associated with feelings of discomfort (Bessenoff & Sherman, 2000; Dasgupta & Rivera, 2006; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Dovidio, Kawakami, & Gaertner, 2002; Fazio Jackson, Dunton, & Williams, 1995; McConnell & Liebold, 2001), concern about how to act or appearing racist (Shelton, 2003; Trawalter, Richeson, & Shelton, 2009) and physiological arousal (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001). Regulating this anxiety distracts individuals by focusing them on the self (e.g., Vorauer, Hunter, Main & Roy, 2000), depletes cognitive resources, and negatively affects performance on subsequent tasks (Govorun, & Payne, 2006; Park, Glaser, & Knowles, 2008; see Richeson & Shelton, 2007 for a review). We hypothesize that greater implicit bias increases whites’ anxiety when teaching black students, and that the resultant distraction and depletion should diminish the quality of their instruction and, subsequently, student learning. Given that white educators teach the vast majority of racial minorities in the U.S. (Howard, 2006) and implicit racial bias is prevalent in modern society (Nosek et al., 2007), the cumulative effect of such a process across teachers and over time could significantly contribute to overall racial disparities in academic achievement.
Recent research generally supports the contention that implicit bias can harm minorities’ academic outcomes. In particular, Van den Bergh and colleagues (2010) showed that the implicit bias of native Dutch teachers was negatively related to the academic performance of students who were Moroccan and Turkish, but not those who were Dutch. However, unlike the present work, the authors hypothesized and found that this effect was partially mediated by teachers’ biased expectations for students. Lower expectations can detract from student learning through a variety of mechanisms including less interpersonal warmth and less effortful teaching (e.g., Brophy, 1986), but there is no reason to believe academic expectations for students should affect instructors’ anxiety. In contrast, we suggest greater implicit bias is associated with greater anxiety and managing this anxiety impairs the quality of teachers’ instruction of Black students.

To test the hypothesis that, controlling for explicit prejudice, white instructors’ implicit racial bias negatively impacts their performance when teaching black students, we simulated teacher-student interactions with cross-race and same-race dyads (Study 1). We predicted that greater implicit bias among white instructors would predict lower test performance by black, but not white, learners, regardless of instructors’ explicit prejudice. We furthermore predicted that this effect would be mediated by greater instructor anxiety and worse quality lessons to black learners. To firmly demonstrate the mediational role of instructor performance in this process, we employed an experimental approach similar to Word, Zanna, and Cooper (1974): New learners watched video of the same lessons given to blacks in Study 1 and were given the same test of the material (Study 2). It was expected that instructors’ implicit bias would once again negatively predict the performance of these new learners, even though they would not be subject to situational identity threats possibly faced by the original black learners.

Study 1
Method

Participants. Two hundred ten Princeton University undergraduates participated in the study for payment ($20) or course credit to produce 51 cross-race and 54 same-race lesson dyads, meeting our target of fifty of each type of dyad. All dyads were same-gender pairs (37% male).

Procedure. One of three white female experimenters welcomed two previously unacquainted participants at staggered times and directed them to separate rooms so they did not meet before the study. The first participant, who was always white, was assigned to the instructor role. The second participant, who was black or white, was assigned to the learner role.

Instructors first completed an ostensible cognitive flexibility task that in reality was a subliminal priming task inconspicuously assessing implicit racial bias. Next, instructors were seated at a table and given the materials for creating their lesson\(^1\) and eighteen minutes to prepare.

After preparation, the participants were introduced, learners were seated across from the instructor, and the lesson commenced. Each lesson lasted seven minutes\(^2\) and was videotaped. After the lesson, participants were given a five-minute discussion period, which was unstructured and could end at any point. Afterward, participants were separated and the learner was given five minutes to complete a test of the lesson. Instructors completed a measure of explicit bias. Both also completed additional measures not analyzed for this paper. (Experimenter scripts, lesson materials, and history test in Supplementary Materials.)

Measures

---

\(^1\) Byzantine history was the chosen subject matter. It pretested as being a relatively unknown topic yet exhibited ecological validity as material that might be presented in a high school or college course.

\(^2\) Pretesting showed that eighteen minutes was enough to prepare the lesson and seven minutes was enough to cover necessary information without leaving unused time.
**Implicit racial bias.** To measure instructors’ implicit racial bias, we used the subliminal priming task developed by Dovidio and colleagues (1997; 2002). Unlike other implicit bias measures it does not appear to be about racial prejudice and does not entail categorizing stimuli by race. Thus, it minimizes the degree to which subsequent behaviors are an artifact of these factors (Fazio & Olson, 2003). Higher scores on the task indicate greater anti-black/pro-white bias. (Details in Supplementary Materials.)

**Explicit racial bias.** Instructors’ explicit racial prejudice were assessed using the twenty-item Attitudes Toward Blacks (ATB; $\alpha=.84$) Scale (Brigham, 1993). Implicit and explicit racial bias were not correlated ($r=.12, p=.211$).

**Instructor behavior.** Videotapes of the cross-race lessons were edited so the learner was not visible and then rated by three coders blind to instructors’ implicit bias. Two coders rated instructor anxiety ($r=.74, p<.001$) on four dimensions using a 7-point Likert scale (i.e., nervousness, discomfort, awkwardness, and stiffness; 1=“not at all”, 7=“very”; $\alpha=.83$). Two coders rated lesson quality ($r=.85, p<.001$) on four dimensions using a 7-point Likert scale (i.e., clarity, ease of following the lesson, monotonous information delivery (R), and amount of effort; 1=“not at all”, 7=“very”; $\alpha=.90$).

**Learner test performance.** Learners’ performance was assessed using thirty-three test questions, each requiring only a few words to answer. The question content was drawn directly from the lesson notes provided to the instructor and included names, dates, and facts (see Supplementary Materials). Two coders, blind to learner race and instructors’ level of implicit bias, independently scored the tests ($r=.99, p<.001$). We computed a final score by averaging the two raters’ scores ($M=12.71, SD=4.46$; 38 possible points).

**Results**
We conducted a linear regression analysis with instructors’ mean-centered implicit racial bias, learner race (0=white; 1=black), and the interactions of implicit bias and learner race as predictors of the learner’s test performance. Given our interest in the effects of instructors’ implicit bias above and beyond their explicit prejudice, we also entered mean-centered explicit bias and its interaction with learner race as covariates in the regression.

**Learner test performance.** There was a significant main effect of learner race ($\beta=-.198$, $t(99)=-2.05$, $p=.043$, $r^2=.04$) that was qualified by a marginally significant interaction of instructor implicit bias and learner race ($\beta=-.238$, $t(99)=-1.77$, $p=.080$, $r^2=.04$). As predicted, greater implicit bias marginally predicted worse learner test performance for black learners ($\beta=-.241$, $t(48)=-1.82$, $p=.075$, $r^2=.08$) but not for white learners ($\beta=.132$, $t(51)=.97$, $p=.338$, $r^2=.02$; Figure 1).

**Instructor behavior.** We next tested the hypothesis that, controlling for explicit prejudice, greater implicit bias predicts greater anxiety, which in turn predicts lower lesson quality and learner test performance. We tested this mediation pathway using Preacher and Hayes’ (2008) PROCESS method with 10,000 bootstrapped samples and 95% confidence interval, the unstandardized effect of implicit bias on instructor anxiety was significant ($B=5.87$, $SE=2.75$, $p=.038$, CI:.33, 11.40). In the next step, the effect of anxiety on lesson quality was significant ($B=-.55$, $SE=.12$, $p<.001$, CI:.80, -.31). In the final step, the effect of lesson quality on learner test performance was significant ($B=.82$, $SE=.39$, $p=.040$, CI:.04, 1.60). Testing the model as a whole, the indirect effect of instructor implicit bias on learner test performance via

---

3 All reported patterns of effects for implicit bias remained the same without controlling for explicit bias.
4 The interaction of instructor explicit bias and learner race was significant, and greater explicit bias significantly predicted worse learner test performance for black, but not white, learners. However, the effect of explicit bias on black learner performance was not mediated via the predicted pathway. Statistics in Supplementary Materials.
instructor anxiety and lesson quality was significant ($B=-2.68, SE=2.41, CI:-10.03, -0.01, r^2=.29$; Figure 2). The direct effect was reduced to non-significance ($B=-7.62, SE=6.53, CI:-20.77, 5.53$), suggesting full mediation.

**Discussion**

Study 1 demonstrated that greater instructor implicit bias predicted lower test performance for black, but not white, learners, controlling for instructor explicit prejudice. Furthermore, our evidence suggests this was due to higher bias instructors being more anxious and therefore giving less effective lessons. Although greater instructor explicit prejudice also predicted worse performance for black, but not white, learners, this effect was not mediated by the same pathway (See Supplementary Materials), suggesting that explicit prejudice affects learning via distinct mechanisms compared to implicit bias.

**Study 2**

We conducted Study 2 to rule out the alternative hypothesis that the effects in Study 1 were driven solely by blacks’ concern about being targets of prejudice or other responses on the part of black learners themselves. If worse lesson quality actually led to decreases in black learners’ performance as hypothesized, a new sample of non-black participants watching these lessons on video should also demonstrate lower test performance.

**Method**

**Participants.** One hundred sixty-five non-black Princeton University undergraduates (98% white; 38% male) participated for course credit, reaching our goal of an average of three participants per lesson video.

**Procedure.** Participants viewed one randomly assigned cross-race lesson from Study 1 that had been edited so the original learner was not visible. One lesson video was lost, leaving
fifty videos used in the current experiment. After watching the video, participants completed the same test administered in Study 1.

Measures

Learner test performance. Two coders, blind to instructors’ level of implicit and explicit bias, scored the tests ($r=.99, p<.001$). We computed a final score by averaging the two raters’ scores ($M=10.06, SD=4.45; 38$ possible points).

Results

We conducted a linear regression analysis with instructors’ implicit racial bias predicting learners’ test performance. Explicit racial bias was included as a covariate.

Learner test performance. As predicted, greater instructor implicit bias predicted lower learner test performance ($\beta=-5.07, t(162)=-2.44, p=.016, r^2=.04$).5

Instructor Behavior. We used the same behavioral ratings from Study 1 to test mediation, with the omission of the missing lesson video. Identical to Study 1, we tested whether, controlling for explicit prejudice, greater implicit bias predicts greater anxiety, which in turn predicts lower lesson quality and learner test performance. In the first step of this mediation pathway, the unstandardized effect of implicit bias on instructor anxiety was significant ($B=5.97, SE=1.748, p<.001, CI:3.04, 8.39$). The effect of anxiety on lesson quality was also significant ($B=-.44, SE=.31, p<.001, CI:-.68, -.42$). Finally, The effect of lesson quality on learner test performance was significant ($B=1.12, SE=.31, p<.001, CI:.51, 1.73$). Retesting the full model, the indirect effect of instructor implicit bias on learner performance response via instructor anxiety and lesson quality remained significant ($B=-3.66, SE=1.51, CI:-7.43, -1.38, r^2=.12$;  

---

5 Instructor explicit bias did not predict learner performance. Statistics in Supplementary Materials.
Figure 3). The direct effect was reduced to non-significance ($B=-9.50$, $SE=5.16$, CI:$-19.68, .69$), suggesting full mediation.

**General Discussion**

The present research showed that when teaching black learners, greater implicit bias among white instructors was associated with more anxiety, which interfered with instructors’ ability to deliver clear, easy-to-follow lessons. This, in turn, hindered student learning as assessed by a test of the material (Study 1). Study 2 showed that the test performance of non-black learners who watched video of the original lessons were similarly affected by instructors’ implicit bias, suggesting that black learners’ performance in Study 1 was not due to alternative psychological factors, including identity threats. These effects were substantial: Lessons delivered in cross-race contexts by implicitly biased instructors (+1 SD) hindered black learners’ performance by 13% compared to instructors one standard deviation below (Study 1) and reduced non-black learners’ performance by 15% (Study 2). By most grading rubrics, this difference corresponds to over a full letter grade. It is also possible these effects are exacerbated when the subject matter is even more complicated or the learning objectives are more abstract than in the present studies.

The present research revealed that instructors’ implicit bias affects their lessons and their students’ subsequent performance irrespective of instructors’ explicit prejudice. However, we found that greater instructor explicit prejudice also predicted lower learner test performance in Study 1, but not in Study 2. This discrepancy across experiments suggests that instructors’ implicit and explicit biases affect student learning in distinct ways. In line with this thinking, instructors’ implicit, but not explicit, bias predicted their anxiety when teaching black learners, replicating prior research that implicit bias uniquely relates to automatic, difficult-to-control
responses. Perhaps explicitly prejudiced instructors exhibited more deliberately prejudicial behaviors that increased identity threat (Steele et al., 2002) or expectancy confirmation processes (Weinstein et al., 2004) with black learners but that did not affect non-black learners who watched these lessons on video.

The present work complements previous research on how identity threat and expectancy confirmation disadvantage black students in academic contexts to highlight how prejudice can produce vicious cycles. When anxiety and poorer lesson quality associated with instructors’ implicit bias causes black students to perform worse, their relatively poor performance may trigger identity threats and belonging concerns that further diminish performance. These consequences then negatively affect instructors’ evaluations of, and expectations for, the students, thereby further hindering student success. This process may even begin with instructors who are not prejudiced, given that concerns about being viewed as such may increase anxiety when interacting with blacks (Carr, Dweck, & Pauker, 2012; Goff, Steele, & Davies, 2008). In this way, seemingly small, initial effects of subtle prejudice may produce iteratively worse outcomes for students, leading to drastically different long-term outcomes than if the effects of prejudice had been addressed (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). Future work should test both the recursive impact of instructors’ implicit bias and whether these effects generalize to other groups, including female, overweight, and low socioeconomic status students.

Although one limitation of this work is the simulated nature of our teacher-student interaction instead of examining these effects in actual classrooms, there are reasons to be confident regarding the generalizability of our findings. Cross-race teaching interactions in which white teachers instruct black students are the norm in the United States (Howard, 2006), and new college graduates with limited training are disproportionately employed as teachers in
schools with large proportions of racial minorities (Peske & Haycock, 2006). The instructors found in many interracial classrooms are therefore remarkably similar to the instructors in our experiments. Nonetheless, given potential differences between teachers and our student sample, as well as the fact that the effects of implicit bias were marginally significant in Study 1, future research should be conducted in which the quality of actual teachers’ lessons and students’ grades are examined. Moreover, examining these processes in the field would afford the opportunity to examine a host of intriguing hypotheses. For example, it may be the case that highly qualified teachers are less susceptible to the demonstrated effects because anxiety is less apt to hinder, and may even enhance, their expert performance (Michaels, Blommel, Brocato, Linkous & Rowe, 1982). Also, the critical role of lesson quality in the present effects suggest that implicitly biased teachers may undermine the learning of all students in mixed ethnicity classroom containing a significant number of Black students.

Overall, the current work adds to a growing body of research showing that implicit racial bias detracts from the success and well-being of racial minorities, and so highlights critical pathways by which inequality can be reduced.

Acknowledgments

Thanks to experimenters Steph Collelo, Maddy Bernstein, and Kelly Hoffman for their considerable efforts, and to Andreana Kendrick for her contributions.
References


Howard, G. R. (2006). *We can't teach what we don't know: White teachers, multiracial schools*. Teachers College Press.


Fig. 1. (Study 1) Effect of instructors’ implicit racial bias on learner test performance, by learner race. + p < .10.
When Bias Binds

Fig. 2. Correlational model with coder-rated instructor anxiety and lesson quality mediating the relationship between implicit racial bias and black learner test performance. All effects control for explicit racial bias. Coefficients are unstandardized. + p<.10, * p<.05, *** p<.001.

Indirect Effect CI: [-10.03, -.01]

Indirect Effect CI: [-7.43, -1.38]
Fig. 3.

(Study 2) Re-tested correlational model with coder-rated instructor anxiety and lesson quality from Study 1 mediating the relationship between implicit racial bias and new non-black learner test performance. All effects control for explicit racial bias. Coefficients are unstandardized. + p < .10, * p < .05, *** p < .001.
Manuscript Highlights

- We simulated learning interactions with white instructors and black or white learners.
- Instructors’ implicit racial bias predicted black, but not white, learner test performance.
- This effect was mediated by instructors’ anxiety and lesson quality as rated by coders.
- Instructors’ implicit bias also predicted test performance of non-black participants who later watched the video recorded cross-race lessons.