

The Mentor's Dilemma: Providing Critical Feedback Across the Racial Divide

Geoffrey L. Cohen
Claude M. Steele
Lee D. Ross
Stanford University

Two studies examined the response of Black and White students to critical feedback presented either alone or buffered with additional information to ameliorate its negative effects. Black students who received unbuffered critical feedback responded less favorably than White students both in ratings of the evaluator's bias and in measures of task motivation. By contrast, when the feedback was accompanied both by an invocation of high standards and by an assurance of the student's capacity to reach those standards, Black students responded as positively as White students and both groups reported enhanced identification with relevant skills and careers. This "wise," two-faceted intervention proved more effective than buffering criticism either with performance praise (Study 1) or with an invocation of high standards alone (Study 2). The role of stigma in mediating responses to critical feedback, and the implications of our results for mentoring and other teacher-student interactions, are explored.

The mentor who wants to provide constructive performance feedback faces a dilemma. Information about shortcomings in the student's work, and pointed suggestions for improvement, can undermine the student's self-confidence and motivation to succeed. Providing critical feedback that encourages rather than discourages the recipient is a challenge for all teachers (Bruner, 1962; Daloz, 1986; Sansone, Sachau, & Weir, 1989), tutors (Lepper, Aspinwall, & Mumme, 1990), managers (Weisinger, 1990), coaches (Horn, 1985), and other educators. But the dilemma is particularly acute when potentially threatening scholastic feedback must be provided to minority students facing negative stereotypes about their group's intellectual capacities.

Minority students are aware that important people in their schooling environment may doubt their ability and belonging (Steele, 1992, 1997; Steele & Aronson, 1995).

Critical feedback, accordingly, may be especially threatening to these students because instead of merely offering information about areas in need of improvement, it raises the prospect that they have been judged in light of a negative stereotype. This "attributional ambiguity" constitutes a double-edged sword (Crocker & Major, 1989; Crocker, Voelkl, Testa, & Major, 1991). Although it may protect students' self-esteem by allowing them to attribute negative feedback to racial bias rather than shortcomings in their own performance, it also may lead them to dismiss rather than act on potentially useful criticism—especially when the criticism comes from a White evaluator rather than a Black one (Banks, Stitt, Curtis, & McQuarter, 1977; Crocker & Major, 1989; Crocker et al., 1991; see also Kleck, 1966).

"Stereotype threat" may create additional consequences for minority students' motivation upon receiving negative feedback (see Steele, 1992, 1997; Steele & Aronson, 1995; Spencer, Steele, & Quinn, 1997). In the context of a broadly disseminated stereotype, the decision to renew or increase one's efforts carries a threatening possibility. While further effort increases chances for success, it also may increase the cost of failure because such failure threatens to confirm the alleged limitation in ability both in the eyes of others and perhaps in one's

Authors' Note: We thank Jennifer Janopaul and Ana-Christina Ramon for their tireless assistance on this project. We also are grateful to Ronaldo Mendoza and Emily Pronin for their constructive comments throughout this project. Funding for this project was provided by a National Research Service Award (NIMH) Pre-Doctoral Fellowship, by National Institutes of Mental Health (Grants MH44321 and MH51977) and Russell Sage Foundation (Grant 879.304). Correspondence concerning this article should be addressed to Geoffrey L. Cohen, who is now at the Department of Psychology, Yale University, P. O. Box 208205, New Haven, CT 06520; e-mail: geoffrey.cohen@yale.edu.

PSPB, Vol. 25 No. 10, October 1999 1302-1318
© 1999 by the Society for Personality and Social Psychology, Inc.

own eyes as well. Rather than expose themselves to such peril, minority students may respond to critical feedback by withdrawing from the performance task, especially insofar as that feedback made them doubt their ability to reach the higher performance standard it demanded. To further protect self-regard, minority students also may adopt the psychological defense of disidentification, that is, diminishing the importance of the relevant domain of achievement as a basis of self-esteem (Steele, 1992, 1997). This self-protective response may further discourage the effort and risk-taking that could boost self-esteem through success in meeting higher performance standards.

The challenge facing the mentor is thus to provide feedback in a *wise* manner, that is, in a manner that discourages expectations and attributions of racial bias and that minimizes stereotype threat. We borrow the term wise from the keen-eyed sociologist Erving Goffman (1963), who in turn borrowed it from the gay subculture of the 1950s. In its original use, the term referred to non-stigmatized individuals who were recognized for their ability to see the full humanity of those bearing a stigma. Our use vis-à-vis intervention strategies has a similar connotation. Wise strategies for assisting minority students are those that assure the students that they will not be judged stereotypically—that their abilities and “belonging” are assumed rather than doubted (Steele, 1992, 1997).

Some features of wise strategies are suggested by successful interventions that, in defiance of troubling statistics on minority achievement, have raised the grades, test scores, and college prospects of at-risk and minority youth (Comer, 1980, 1988; Kleinfeld, 1975; Lepper et al., 1990; Lepper, Woolverton, Mumme, & Gurtner, 1993; Mathews, 1988; Treisman, 1985; see Cose, 1997; Steele, 1997, for reviews). The educators in these programs all refute negative stereotypes by conveying a clear faith in each student's intellectual potential. But they do not impart this message by assigning easier work to ensure student success or by offering heavy doses of unstinting praise—all too common tactics of well-meaning but unwise teachers. Indeed, several authors offer detailed discussions of the dangers of “overpraising” and “under-challenging” members of stereotyped groups (Dweck, Davidson, Nelson, & Enna, 1978; Fernandez, Espinosa, & Dornbusch, 1975; Harber, 1998; Kleinfeld, 1975; Massey, Scott, & Dornbusch, 1975; Mathews, 1988; Steele, 1992, 1997; see also Barker & Graham, 1987; Dawes, 1994; Meyer et al., 1979; Mueller & Dweck, 1998). Rather, minority students in all of these otherwise diverse success stories are challenged with high performance standards—standards that *presume* their motivation and ability to succeed. The educators often go an important step

further by explicitly assuring students of their capacity to meet those standards through greater effort.

Anecdotal evidence illustrates the effectiveness of setting and maintaining high standards. Jaime Escalante (whose work was portrayed in the movie *Stand and Deliver* and documented by Mathews, 1988) challenged his East Los Angeles Latino students to take and pass the advanced placement (AP) exam in calculus. Escalante's students met this standard. In fact, they accounted for 27% of all Mexican Americans receiving college credit on their AP exam, and the rate of advanced placement compared favorably with that obtained in many privileged, suburban schools. Xavier University, which, despite its small size and scant endowment, sends more Black students to medical schools than any other university, and Georgia Tech, which enjoys exceptional success in graduating minority students from its engineering curriculum, similarly set highly demanding standards (see Steele, 1992, 1997; also Cose, 1997).

The invocation of high standards is apt to be of limited value unless the student is assured, implicitly or explicitly, that he or she is capable of reaching the higher standard. Successful interventions thus continually convey the message that students can succeed through effort and persistence (Mathews, 1988; see also Cose, 1997). In a sense, the message is that academic ability, or even so-called intelligence, is not fixed or immutable. Rather, it can be enhanced through effortful practice and the cultivation of specific skills (see Dweck, Chiu, & Hong, 1995; Dweck & Leggett, 1988; Ericsson, 1993). As Dweck and her colleagues have shown (see also Nicholls, 1978), students who believe in the immutability of intelligence focus on “performance goals”; they seek to demonstrate rather than enhance their competence and are apt to withdraw from tasks where they risk failure. In the face of the inevitable setbacks and frustrations that accompany the pursuit of high standards, stereotype-threatened students may thus need to be assured that, with greater effort, they can increase their capacities.

Not coincidentally, the most successful academic programs aimed at minority youth often are presented as honorific rather than remedial, with correspondingly higher standards for student performance (Steele, 1992, 1997). The assurance that students can reach the higher standards is implicit in their recruitment. Students in Urie Treisman's (1985) rigorous calculus workshops at Berkeley, for example, are told that their invitation to join the “special honors program” is based on their prior records and demonstrated academic potential. Other features of the program reinforce high standards and refute alleged limitations in ability. For example, interracial group study sessions allow students to observe one another wrestle with the assigned work. They come to see their own frustration with the course material not as

unique to themselves (and thus an indication of limited ability) but as a response to demanding work that is common to all students, Black and White alike (see also Prentice & Miller, 1993). Treisman's program raised the percentage of Black undergraduates who passed 1st-year calculus from 60% to 97%. In spite of the program's small size—students meet 6 to 8 hours each week during their freshman year—it also increased the percentage of Black students who ultimately completed college by 40%.

Our review suggests the wisdom of combining the invocation of high standards with the assurance of students' capacity to reach those standards. Such a two-faceted intervention, we argue, is especially germane to the concerns of the mentor who must provide feedback across racial lines. The invocation of high standards alone may encourage minority students to attribute critical feedback to rigorous performance standards rather than to racial bias (see Banks et al., 1977). But the assurance that the student can meet the higher standard may be necessary to diminish the threatening possibility that the student has reached a group-based limitation in ability. In a sense, it is the explicitness of these messages that may be disproportionately important for students who face negative stereotypes. Nonminority students receiving unbuffered rigorous criticism, we argue, will be more inclined than minority students to automatically infer high standards and to further assume that the critic felt that they could meet these standards.

The two experiments reported here examined the response of Black and White students to critical feedback. Our first study pitted the effectiveness of the wise combination of explicit invocation of high standards, and assurance that the student could meet those standards, against the more commonplace tactic of accompanying criticism with a buffer of performance praise. Providing such a buffer of praise is a strategy advocated both by some organizational theorists and some educators of minority students and one frequently practiced by White evaluators giving feedback to Black students (Harber, 1998). Our second study compared the effects of the same wise, two-faceted intervention with those of invoking high standards in the absence of the relevant personal assurance.

The paradigm we developed was one in which students wrote and then received critical feedback on a letter of commendation for their most memorable and effective teacher, coach, or mentor—with the prospect of ultimately having their letter published. This paradigm was noteworthy in several respects. First, in contrast to the performance task used in many feedback studies, our students' task was highly motivating, personally relevant, and one for which they were uniquely qualified. Second, the relevant feedback consisted not of an arbitrary numerical ranking or rating of performance

but of detailed criticisms and suggestions for improvement. Students could see a clear (indeed, authentic) connection between the feedback and the specifics of their work. Third, our dependent variable measures included particularly natural and appropriate measures of task motivation, that is, students' expressed belief in their capacity to improve their letter and their interest in undertaking further revision.

STUDY 1

In the first of two experimental sessions, Black and White college students were asked to write a letter of commendation for their favorite teacher, ostensibly for potential publication in an education journal. In the second session, a week later, students in three conditions received critical feedback pointing out weaknesses and suggesting strategies for improvement from a reviewer that they presumed to be White. In an "unbuffered criticism" condition, the students received this critical feedback with no further comment. In a "wise criticism" condition, students received the same critical feedback buffered with an explicit invocation of high standards and an assurance of the particular student's capacity to reach those standards. In a "positive buffer" condition, students received the critical feedback buffered only by general praise of their performance.¹

Three types of dependent variable measures were used. One measure assessed students' perceptions of the bias showed toward them on the part of the reviewer. A second set of measures assessed motivation at the task, specifically, participants' belief in their ability to improve their work with greater effort and their interest in revising and resubmitting their letter. (Such measures were specifically attuned to the concerns of the mentor, that is, whether the student, upon receiving critical feedback, believes that he or she can now do better and wishes to try again.) A third set of dependent measures assessed general feelings of identification with the more global academic skills relevant to the letter-writing task. (In addition to their intuitive appeal, the validity of our motivation and identification dependent measures is supported by a principal component analysis reported in each of the Results sections.)

Our working hypothesis (following Crocker et al., 1991, and Banks et al., 1977) was that Black students would respond to unbuffered criticism more negatively than would White students. That is, Black students would rate the criticism as more biased, and feel less motivated, than their White peers. Our main concern, however, was the extent to which the tactic of accompanying the criticism with an explicit invocation of high standards and an assurance of the student's capacity to reach those standards would attenuate those negative effects and thus reduce or eliminate the relevant race differences in

response to criticism. We further predicted that our two-faceted intervention would reduce the differences between Black and White students' responses to criticism more than would the inclusion of general praise to buffer the criticism.

METHOD

Participants

All but 4 participants were recruited by telephone from a registrar's list of Stanford undergraduates and paid \$10 for their participation. Only self-classified Black and White students who had not taken any courses in introductory or social psychology (and thus were unlikely to be suspicious about the purpose of the study) qualified for participation in the study from this list. The other 4 participants (2 Blacks and 2 Whites) were recruited by telephone from an introductory psychology course within the 1st week of the course before any discussion of relevant issues and subject matter; these students received course credit for their participation.

A total of 45 Black and 48 White students were assigned at random to the three criticism conditions. However, data from 2 of these participants were discarded prior to analysis, 1 Black student who voiced suspicion about our interest in race and 1 White student who received feedback that erroneously referred to his male teacher as "she."

Procedure

Students participated in the study individually, although in some instances, 2 students heard the initial introduction to the study together before being escorted into separate rooms to complete the experimental task.² Each student took part in two sessions, writing an account of a favorite teacher during the first session and receiving feedback on that account during the second session.

Session 1. Each student was greeted by an experimenter (a White male or White female) who escorted him or her to a laboratory room and then proceeded to elaborate on the ostensible purpose of the study. Students were told that the study concerned the psychology of effective teaching and that its objective was to uncover those qualities that define excellence in teaching. Their task, it was explained, would be to write a "letter of commendation" for their favorite teacher, who could be an elementary or secondary schoolteacher, a university professor, a tutor, or a coach—anyone who they felt had "made an important contribution to their personal growth." Students were led to believe that the researchers would seek to publish the best letters in an education journal.

To strengthen the link between the letter-writing task and more general academic skills relevant to negative minority-group stereotypes, students were told that their task would "require the same skills necessary for writing an effective paper. . . . You'll have to communicate your ideas clearly, make intelligent points and back them up with supporting arguments, and maintain a coherent organization and structure." The experimenter went on to explain that in the first session, students would write the letter and that in the second session a week hence, their letters would be returned to them with comments provided by a reviewer from a research panel—comments that might help them improve their letter if they decided to revise it.

Students next were asked to read and sign a participation consent form. As they were reading the form, the experimenter reminded participants of a request (made at the time of their recruitment) that they bring to the study a photograph of themselves, ideally with their favorite teacher, that would accompany their letter should it be published. When (in all but two cases) students indicated that they had forgotten or been unable to find such a photograph, the experimenter said, "No problem, but if it's okay with you, we'd like to take your picture at the end of today's session because we like to get snapshots of our recommenders for the journal in case we decide to publish what you write." All students agreed to this request. (In the two cases in which the student had brought a photograph, the experimenter said, "Good. We'll attach that to your letter at the end of today's session because we like to get snapshots of our recommenders for the journal in case we decide to publish what you write.") This use of the photographs would later serve to alert students that their racial identity would be known to their evaluator.

Letter-writing task. Printed across the top of the letterhead that students were told to use for their task was the heading "Letter of Recommendation for _____ for Recognition in Teaching Excellence." Students were instructed to write the name of their teacher on the blank line of this header, to begin writing their letter on the same page, and to use additional blank sheets as necessary. They also were told that they could write a rough draft or outline first if they liked. After 40 minutes, the experimenter checked on the student's progress. Students were given as much time as they needed to finish their letters, and most finished within 45 minutes (with a range from 20 to 75 minutes).

Session 1 (pre-manipulation) measures. After completing the letter-writing task, all students were given an initial questionnaire that included three critical items designed to assess motivation regarding the letter-writing task. These items asked, "How much do you think

you could improve your letter given the opportunity to revise it?" "How much interest do you have in revising your letter before submitting it to the journal for the final time?" and "How important is it to you to do a good job on your letter of recommendation?"³ Three other items measured identification with the more global skills underlying both the specific letter-writing task and written communication more generally. These items asked, "How important is being a good writer to your view of yourself?" "How open are you to career options that would require writing skills?" and "How would you rate your overall competence as a writer?" Students responded using appropriately labeled 7-point scales.⁴

After completing the questionnaire, students were photographed using a Polaroid Instant Camera. The experimenter explained again that a copy of the photo would be used in the journal should their entry be accepted for publication. Participants were then asked to staple their photograph to their letter, thereby ensuring that they were aware that their race would be known by the reviewer who would read and evaluate their letter. They also were asked to put their entry in a large mailing envelope to be deposited in a box in the hallway upon their departure. Before leaving, students were scheduled for the next session, given a reminder slip listing its date and time, and either paid \$5 or assured that they would receive the relevant course credit.

Session 2. Approximately 1 week later, students returned for the second session, whereupon the experimenter greeted them and handed students a sealed envelope containing both their letter and a handwritten review by an alleged reviewer. Students were told that they would be allowed to revise their letter and resubmit it to the journal. The experimenter told each student to take his or her time reading their letter and the review, then left the room. The first page of this review was a summary sheet that indicated the student's name and entry number, their teacher's name, and the name of the purported reviewer who had evaluated their letter (i.e., "Dr. Gardiner Lindsay," which is a name that pilot testing had established as recognizably Caucasian).

Critical feedback. Each letter was marked with handwritten corrections of errors in spelling and grammar; in addition, there were comments (approximately eight in number) regarding particular shortcomings in style, structure, and wording (i.e., "awkward," "confusing," or "unclear") as well as specific suggestions for improvement. Each letter also received two check marks, placed in the margins, acknowledging good points. This feedback had been provided by assistants trained in providing useful critical reactions but blind to the letter writer's race and experimental condition.

At the end of each letter, there was a general critique of the student's performance, which was again handwritten and signed by Gardiner Lindsay; the critique, which follows, appeared to be individualized but was actually identical (except in the use of appropriate masculine versus feminine pronouns) for all students receiving critical feedback (and that addressed limitations that were, in fact, characteristic of virtually all the students' letters):

Student's name:

Your letter needs work in several areas before it can be considered for publication. In addition to some routine editorial suggestions that I've offered, most of my comments center on how you could breathe more life into your letter and make the description of your favorite teacher and her [his] merits more vivid, personal, and persuasive. As it stands, your letter is vague and rambling—long on adjectives and short on specific illustrations. You describe your teacher's dedication and commitment but you haven't explained why your teacher is more exemplary in her [his] contribution, more deserving of recognition, than most of the other nominees cited by other writers. In particular, it would be helpful to be more specific when you describe your teacher, to pay closer attention to the details that inspired your high opinion of her [him]. What were some of the specific things your teacher did that set her [him] apart from all other teachers you've encountered in your life? You cover this at certain points in your letter, and it is there that your letter begins to come to life—you need to sustain this.

One last comment: If you choose to revise your letter, you should spend significantly more time explaining your teacher's impact on your own personal growth. What made her [his] influence so much more important than other teachers in your life? Perhaps your teacher opened your eyes to something you hadn't seen before, perhaps she [he] helped you to see your potential. Sometimes you touch on this but you fail to build on it. You need to discuss the long-term imprint [teacher's name] has left on you in greater detail—this enduring impact is perhaps the strongest testimony of a teacher's success.

Experimental conditions. Students returning for the second session were randomly assigned to one of three criticism conditions. Students in the unbuffered criticism condition received the criticism described above without further comment. Students in the wise criticism condition received that criticism prefaced with the following comments designed to ward off potentially destructive attributions by explicitly invoking high standards while assuring the particular student that he or she could meet such standards:

It's obvious to me that you've taken your task seriously and I'm going to do likewise by giving you some straightforward, honest feedback. The letter itself is okay as far as it goes—you've followed the instructions, listed your teacher's merits, given evidence in support of them, and importantly, produced an articulate letter. On the other hand, judged by a higher standard, the one that really counts, that is, whether your letter will be publishable in our journal, I have serious reservations. The comments I provide in the following pages are quite critical but I hope helpful. Remember, I wouldn't go to the trouble of giving you this feedback if I didn't think, based on what I've read in your letter, that you are capable of meeting the higher standard I mentioned.

By contrast, students in the positive buffer condition received the same general criticism described above but prefaced with the following comments:

Overall, nice job. Your enthusiasm for your teacher really shows through, and it's clear that you must have valued her [him] a great deal. You have some interesting ideas in your letter and make some good points. In the pages that follow, I've provided some more specific feedback and suggested several areas that could be improved.

Session 2 (post-manipulation) measures. After reading their letter and review, students filled out a questionnaire containing the same task motivation and identification items that had been included in the Session 1 questionnaire. An additional item, embedded among three other incidental items on a page titled "Impressions of the Reviewer," asked students to rate the reviewer's bias on a scale from 1 (*not at all biased*) to 7 (*extremely biased*).

Debriefing. At the conclusion of the second session, in a postexperimental interview, students were first probed for possible suspicion and then informed about the general purpose of the study. The experimenter not only revealed the deceptive elements in the feedback and purported rationale for the study but also offered an account of the study's real concerns and the psychological processes under investigation (using a "process debriefing" procedure espoused by Ross, Lepper, & Hubbard, 1975). Students were informed of the study's focus on strategies for attenuating adverse responses to criticism, but most were not informed of our interest in race because such a disclosure would rule them out as prospective participants for future research on stereotype threat at Stanford and was unnecessary in making their participation in the study an interesting and satisfying experience. Finally, students were thanked for their

participation and either paid \$5 or again assured that they would receive the relevant course credit.

RESULTS

Analyses revealed no main effects or interactions involving gender of participant or experimenter; these variables, accordingly, receive no further consideration in our report.

Creation of "Task Motivation" and "Identification With Writing Skills" Composites

To lend empirical support to the validity of the motivation and identification constructs, a principal component analysis using varimax rotation was performed on the six premanipulation measures of task motivation and identification. As expected, the component analysis produced two discrete components, accounting for 64.7% of the variance in the response measures. The first component, accounting for 38.5% of the variance, encompassed the three items pertaining to identification with writing skills (eigenvalue = 2.31). The second component, accounting for 26.5% of the variance, encompassed the three items pertaining to task motivation (eigenvalue = 1.56). All items loaded highly on their respective components (loadings ranged from .68 to .84) and none had a substantial loading (i.e., greater than .25) on a component other than its own. Premanipulation and postmanipulation composites were created for both components. That is, the three premanipulation items for each component and the three postmanipulation items for each component (standardized to equate their variance) were averaged to provide premanipulation and postmanipulation scores of task motivation and identification.

Ratings of Bias

Preliminary overall analysis. A 2×3 (i.e., Race \times Criticism Condition) ANOVA was performed for ratings of bias. The omnibus ANOVA tests revealed a marginal main effect of race, with Black students rating the reviewer as somewhat more biased ($M = 2.75$) than White students ($M = 2.26$), $F(1, 85) = 3.29, p = .07$, and a borderline Race \times Condition interaction, $F(2, 85) = 3.08, p = .05$, indicating that this race effect varied as a function of criticism condition. Figure 1 displays the six relevant cell means.

Unbuffered criticism versus wise criticism. More important than these overall analyses are the specific comparisons involving Black and White students in the two conditions most relevant to our hypotheses. Our first concern was whether, as predicted, Black students receiving unbuffered criticism would rate the reviewer as more biased

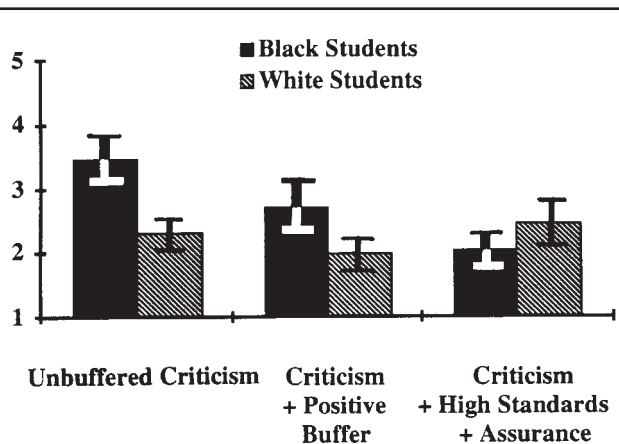


Figure 1 Ratings of bias as a function of race and feedback condition in Study 1.

than would White students. Our second and more critical question was the extent to which our wise intervention, that is, the invocation of high standards accompanied by the assurance that the particular student could meet those standards, would attenuate the adverse effects of such criticism and reduce the relevant race difference.

As Figure 1 attests, Black students receiving unbuffered criticism rated the reviewer as more biased ($M = 3.47$) than did White students receiving the same criticism ($M = 2.31$), $t(85) = 2.53$, $p < .02$. However, as predicted, this between-race difference was eliminated when the criticism was accompanied by the combination of high standards and assurance. Indeed, in the face of such wise feedback, Black students rated the reviewer somewhat lower in bias ($M = 2.04$) than did White students ($M = 2.47$). Evidence for the relevant interaction effect is provided by a weighted contrast comparing the size of the race effect in the unbuffered criticism condition with that in the wise criticism condition (with weights of 0 assigned both to White and Black participants in the criticism plus positive buffer condition). This analysis yielded a significant contrast, $t(85) = 2.41$, $p < .02$. (In the case of ratings of bias, however, between-race comparisons may be inappropriate because White students may have interpreted the term *bias* to mean hostility toward them personally rather than animus toward members of their race. Thus, it is perhaps more relevant simply to note that, as predicted, Black students in the unbuffered criticism condition rated the reviewer as significantly more biased than did Black students in the wise criticism condition, $t[85] = 3.02$, $p < .005$.)

Criticism plus positive buffer condition. In the criticism plus positive buffer condition, the relevant race difference was smaller but still apparent, $t(85) = 1.60$, $p < .12$,

with the mean bias rating provided by Black students ($M = 2.73$) falling squarely between those provided by Black students in the other two criticism conditions (see Figure 1). (A simple within-race contrast, in which Black students in the positive buffer condition are compared with Black students in the wise criticism condition, yields a similar t value, $t[85] = 1.47$, $p < .15$.) A reasonable description of the data—and one that is consistent with our predictions and conceptual analyses—is that the Black-White difference is greatest in face of unbuffered criticism, smaller when the same criticism is buffered by positive feedback, and nonexistent (indeed, slightly reversed) when this criticism is buffered by high standards and a personal assurance. A contrast testing the relevant linear decrease in this racial difference (i.e., assigning weights of +1 to White students in each of the three conditions and weights of -3, -1, and +1 to Black students in the unbuffered criticism, criticism plus positive buffer, and wise criticism conditions, respectively) yields a highly significant t value, $t(85) = -3.551$, $p < .001$, and most important, such a contrast accounts for the vast majority (90%) of between-cell variance. Simply comparing the size of the race effect in the wise criticism condition with that in the other two criticism conditions similarly yielded a significant result, $t(85) = 2.40$, $p < .02$, as did simply comparing the bias rating of Black students in the unbuffered criticism condition with that of Black students in the two other criticism conditions, $t(85) = 2.67$, $p < .01$.

Task Motivation

Preliminary overall analyses. A 2×3 ANCOVA was performed on self-reported, postmanipulation task motivation using premanipulation task motivation as a covariate. This analysis revealed a marginal main effect of condition, such that motivation was lower in the unbuffered criticism condition ($M = -0.20$) than either in the criticism plus positive buffer condition ($M = 0.07$) or in the wise criticism condition ($M = 0.12$), $F(2, 84) = 2.74$, $p = .07$. This overall analysis also revealed a Race \times Condition interaction, $F(2, 84) = 3.11$, $p < .05$. Once again, however, it is more specific comparisons that are of greatest interest (see Figure 2 for the adjusted cell means).

Unbuffered criticism versus wise criticism. As in the case of our bias measure, we expected to find a race difference in motivation in the unbuffered criticism condition but to find such a difference attenuated or eliminated in the two-faceted wise criticism condition. This expectation, bolstered by the finding of a significant association between perceived bias and lack of motivation ($r = -.27$, $p < .01$), proved accurate. In the face of unbuffered criticism, Black students reported lower motivation ($M = -.48$) than did White students ($M = .09$), $t(84) = -2.75$, $p < .01$. By contrast, when the same criticism was

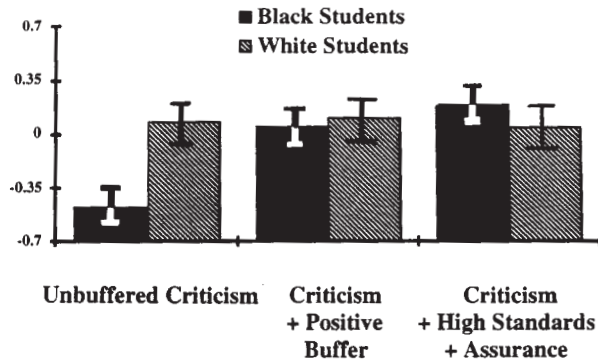


Figure 2 Task motivation as a function of race and feedback condition in Study 1.

accompanied by the invocation of high standards and appropriate assurance, the task motivation of Black students was notably increased. Indeed, they reported slightly higher motivation ($M = .20$) than did White students in the same condition ($M = .05$). A weighted contrast comparing the size of the race effect in the unbuffered criticism condition with that in the wise criticism condition (with weights of 0 assigned both to White and Black students in the criticism plus positive-buffer condition) again yielded a significant contrast, $t(84) = 2.43, p < .02$.

Criticism plus positive buffer condition. Once more, the mean for Black students in the criticism plus positive buffer condition ($M = .05$) fell between those for Black students in the other two criticism conditions, although closer to that in the wise criticism condition than the unbuffered criticism condition. Also, as with ratings of bias, the race difference in responses on the task motivation measure was greatest in the unbuffered criticism condition, smaller in the positive buffer condition, and nonexistent (indeed, slightly reversed) in the wise criticism condition. Again, as with the bias measure, a contrast testing the relevant linear decrease in the racial difference proved highly significant, $t(84) = 3.31, p < .002$, and accounted for 84% of the between-cell variance. Simply comparing the size of the race effect in the wise criticism condition with that in the other two conditions yields a marginally significant result, $t(84) = 1.80, p < .08$.

Identification With Writing Skills

Preliminary overall analysis. A 2×3 ANCOVA was performed for postmanipulation identification, with premanipulation identification as a covariate. The omnibus tests revealed only a marginal main effect of condition, reflecting the fact that among Black and White students alike, identification with writing skills was lower in the face of unbuffered criticism ($M = -.10$), and also in the

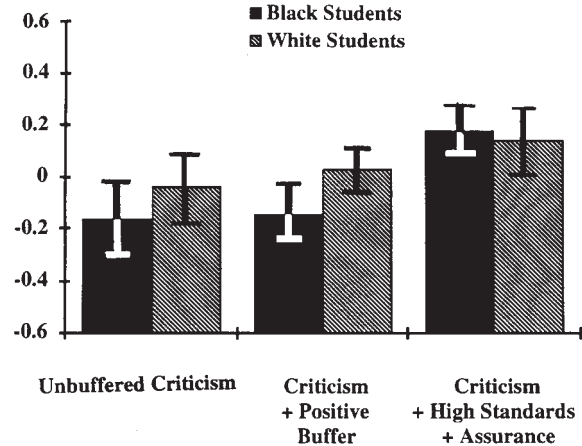


Figure 3 Identification with writing skills as a function of race and feedback condition in Study 1.

face of the same criticism buffered by positive feedback ($M = -.06$), than in the wise criticism condition offering both the high standards invocation and the appropriate assurance ($M = .16$), $F(2, 84) = 2.63, p < .08$. Figure 3 displays the relevant adjusted means.

Unbuffered criticism versus wise criticism. In contrast to the case for the perceived bias and task motivation measures, Black and White students in the unbuffered criticism condition differed little on this measure, $t < 1$, and the weighted contrast involving race and criticism condition failed to show any interaction effect, $t < 1$. This absence of an interaction effect, however, should not obscure the fact that the wise criticism condition accomplished its pedagogical goal. That is, Black students who received wise criticism showed greater identification with writing skills ($M = .18$) than those who received the unbuffered criticism ($M = -.16$), $t(84) = 1.96, p = .05$. The failure to obtain a significant interaction effect simply reflects the fact that White students showed a similar, albeit smaller (statistically insignificant), difference in responding to the two relevant criticism conditions.

Criticism plus positive buffer condition. The mean identification score among Black students in the criticism plus positive buffer condition was as low ($M = -.15$) as that of students receiving unbuffered criticism ($M = -.16$). Simply contrasting these two conditions with the wise criticism condition yielded a statistically significant result, $t(84) = 2.20, p < .05$.

DISCUSSION

As predicted, Black students generally responded less favorably to unbuffered criticism than did White students, both in ratings of bias and in measures of task motivation. The relevant race difference in response was

somewhat reduced when the criticism was buffered with praise of the student's work, but it was totally eliminated (indeed, even slightly reversed) only when the criticism was buffered with the wise combination of an explicit invocation of high standards and an appropriate assurance that the student in question could meet such standards. Our wise criticism condition also best served to bolster the willingness of students, Black and White alike, to identify themselves with the more global academic skills involved in their letter-writing task.

An obvious question arising from this study is whether the effects obtained in the wise criticism condition were due solely to the attributional benefit of evoking high standards or whether the additional assurance regarding the student's capacity to reach those standards was necessary. Study 2 began the effort to disentangle these two elements of the wise criticism manipulation by introducing a condition that invoked high standards but that made no assurance regarding the student's capacity to reach those standards. We predicted that this condition would be less effective than the two-faceted wise criticism condition offering the extra assurance of the student's ability to reach the standard. Specifically, our working hypothesis was that invoking high standards alone might suffice to deflect attributions of bias but would leave intact the stereotype threat created by the delivery of criticism and hence would not enhance task motivation.

The prediction that the simple evocation of high standards would ward off attributions of bias is consistent not only with our conceptual analysis but also with prior findings from the previously cited study by Banks and colleagues (1977). Black participants in that study discounted the objectivity of performance feedback from a White evaluator, relative to a Black one, and also chose to persevere on their own strategies rather than adopt those recommendations made by the White evaluator. Interestingly, this reluctance to trust and to comply with the suggestions of the White evaluator was eliminated when, prior to the exercise, students were told that the evaluator would win money if participants excelled at the task. Simply making explicit the evaluator's goals, it seems, clarified the motives behind the feedback and forestalled attributions of prejudice.

We expected the simpler tactic of making explicit high standards to similarly forestall perceptions of prejudice and racial bias by providing a compelling external attribution for the feedback. That is, in the face of an invocation of high standards alone, Black students would see the criticism as motivated by the reviewer's demands for excellence rather than by any personal or group animus (an attribution perhaps made implicitly by White students in the absence of the explicit invocation of standards). At the same time, we expected this "high standards only" condition to leave unaddressed

the stereotype threat to Black students introduced by the negative feedback. To prevent depressed task motivation, according to our working hypothesis, it would be necessary to attenuate stereotype threat through the additional assurance that the student was personally capable of meeting the higher standard—that is, by providing the same wise two-faceted intervention featured in Study 1.

STUDY 2

The explicit invocation of high standards, accompanied by the assurance of the student's personal capacity to reach those standards, we have argued, is analogous to the strategy used by mentors, teachers, and academic intervention programs that provide critical feedback while successfully addressing the concerns of minority students. The results of Study 1 attested to the efficacy of such feedback. Study 2 attempts to demonstrate that personal assurance is a necessary component of wise critical feedback—that, although the invocation of high standards alone may forestall attributions of racial bias, it does not sufficiently address the negative motivational consequences of the stereotype threat evoked by pointed criticism. (The issue of whether such assurance is not only necessary but sufficient is not addressed by our research design but will be discussed later in this article.) Accordingly, Black and White students in two conditions of our second study again received either critical feedback alone or such feedback accompanied by an invocation of high standards and a personal assurance of their capacity to reach those standards. In an additional condition, however, students received criticism accompanied by an invocation of high standards without such assurance.

Design and Participants

The design of the experiment featured a 2×3 (i.e., Race \times Criticism Condition) factorial design in which 80 Black and 73 White participants (all recruited from a registrar's list) were randomly assigned to condition and received either unbuffered criticism, criticism buffered by an invocation of high standards alone, or criticism buffered both by that invocation and by an assurance of the student's capacity to reach the relevant standards. As in Study 1, ratings of the reviewer's bias, task motivation, and identification with writing skills were used as the dependent measures. The data of 5 participants who either failed to return to their scheduled second session or expressed suspicion about the purpose of the study or the authenticity of the feedback were discarded prior to analysis.

Procedure

Two minor changes were made in the procedure of Study 2. First, students were paid \$12 (rather than \$10) for their participation. Second, the experimenter (again, either a White male or a White female) waited 45 minutes (instead of 40) before checking on the student's progress during Session 1 (to give students more time to produce a satisfactory letter of commendation). The dependent variable questionnaire also was altered slightly. The task motivation composite now included only two items, one asking the extent to which students felt they could improve their performance with further effort and the other assessing students' interest in revising the letter. The writing-skills identification composite now also included two sets of items, one asking the general importance of writing skills to students' view of themselves and a second set of items assessing students' openness to various career possibilities (such as author, journalist, and editor) that demand writing skills (Stoutemeyer & Steele, 1997).

The major change in procedure, however, was the addition of the new high standards only condition and the elimination of the positive buffer (and bland praise) conditions. The unbuffered criticism and wise criticism conditions were identical to those in Study 1. In the new high standards only condition, the criticism delivered to students was prefaced by comments that invoked the relevant standards but then proceeded to justify those standards not with an assurance regarding the student's capacity but with an emphasis on the reviewer's commitment to the journal considering the student's efforts. The comments for the high standards only condition were as follows:

It's obvious to me that you've taken your task seriously and I'm going to do likewise by giving you some straightforward, honest feedback. The letter itself is okay as far as it goes—you've followed the instructions, listed your teacher's merits, given evidence in support of them, and importantly, produced an articulate letter. On the other hand, judged by a higher standard, the one that really counts, that is, whether your letter will be publishable in our journal, I have serious reservations. The comments I provide in the following pages are quite critical but I hope helpful. Remember, I wouldn't go to the trouble of giving you this feedback if I weren't committed to the quality of this journal—I want to uphold the highest standards for what I consider a suitable entry, for you or any student whose work is under consideration.

As noted, all participants again received critical comments and suggestions appropriate to their particular letters along with two critical paragraphs, at the end of the letters, identical for participants in all conditions. After reading the feedback, students completed the

dependent measures. Then, as in Study 1, they were thoroughly debriefed using a process debriefing procedure (Ross et al., 1975), thanked for their participation, and paid.

RESULTS

For all but one measure, there were neither main effects nor interactions involving gender of participant (nor were there effects involving gender of experimenter). Accordingly, with the exception to be noted, this factor will receive no further attention in our report of the results of Study 2. Also, some students omitted answers to one or more questions, resulting in variable degrees of freedom for different measures.⁵

Creation of Task Motivation and Identification With Writing Skills Composites

Principal component analysis, using varimax rotation, again produced two discrete components that accounted for 71% of the variance in the Session 1 premanipulation response measures. The first component, encompassing the two items pertaining to task motivation, accounted for 41.1% of the variance (eigenvalue = 1.65). The second component, encompassing the two items pertaining to identification with writing skills, accounted for 29.4% of the variance (eigenvalue = 1.18). Once again, all items loaded highly on their respective components (loadings ranged from .82 to .85) and none had a substantial loading (i.e., greater than .11) on a component other than its own. As in Study 1, "premanipulation" and "postmanipulation" composites were created both for task motivation and identification, with responses on the two relevant items in each composite standardized to equate their variance and then averaged to provide single premanipulation and postmanipulation scores.

Ratings of Bias

Our basic prediction was that Black students would perceive the reviewer as more biased than would White students in the unbuffered criticism condition but that this difference would be attenuated or even disappear in the two conditions where the criticism was accompanied by an invocation of high standards. A 2×3 ANOVA (i.e., Race \times Criticism Condition) performed on students' ratings of bias (see Figure 4) yielded the predicted Race \times Condition interaction, $F(2, 140) = 3.39, p < .05$.

As in Study 1, Black students in the unbuffered criticism condition rated the reviewer as more biased ($M = 3.04$) than did White students ($M = 2.00$), $t(140) = 2.44, p < .02$. By contrast, Black students in the two high standards conditions rated the reviewer as no higher in bias ($M = 2.30$) than did their White peers ($M = 2.59$). The relevant test of our hypothesis—a weighted contrast

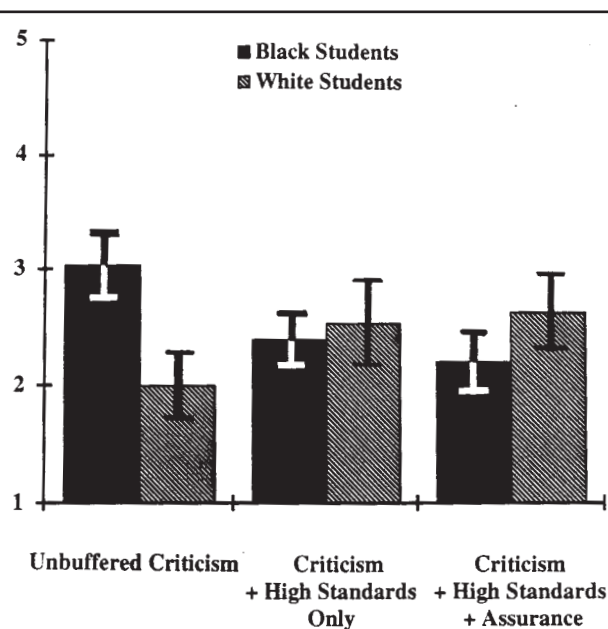


Figure 4 Ratings of bias as a function of race and feedback condition in Study 2.

comparing the size of the race effect in the unbuffered criticism condition with that in the two conditions where the criticism was accompanied by an invocation of high standards—also proved significant, $t(140) = 2.56, p < .02$. (As we noted in reporting the results of Study 1, comparisons of bias ratings by Black and White students may be misleading. However, orthogonal within-race comparisons reveal, as predicted, that Black students rated the reviewer as more biased in the unbuffered criticism condition than in the two high standards conditions, $t[140] = 2.09, p < .05$, whereas there was no reliable difference between their bias ratings in the two high standards conditions, $t < 1$.)

Task Motivation

A simple between-participants analysis of variance was performed on postmanipulation task motivation due to a violation of the homogeneity of regression slopes assumption of covariance analysis. The invocation of high standards alone sufficed to deflect attributions of bias, but it did not prove sufficient to raise Black students' motivation to the level observed in the wise criticism condition featuring both high standards and the added assurance. The relevant analysis yielded a significant Race \times Condition interaction, $F(2, 140) = 3.12, p < .05$. Black students reported significantly greater motivation in the wise criticism condition than they did in either the unbuffered criticism condition, $t(140) = 2.01, p < .05$, or the criticism plus high standards only condition, $t(140) = 2.16, p < .05$. By contrast, the motivation of

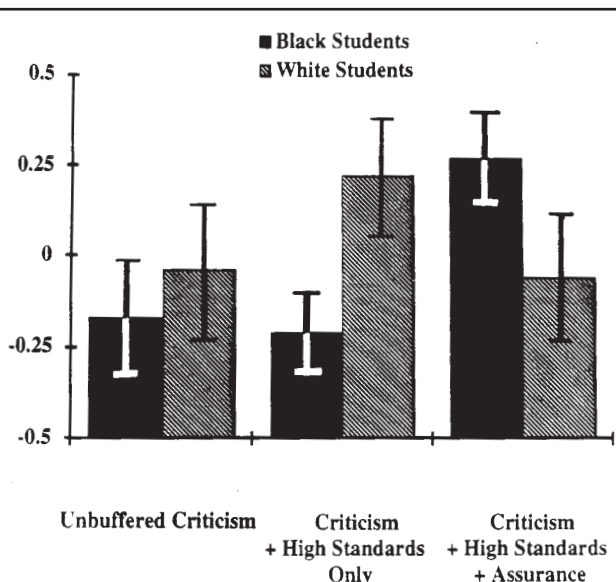


Figure 5 Task motivation as a function of race and feedback condition in Study 2.

White students showed no significant differences across conditions (all relevant contrasts yielded p values $> .19$). Figure 5 displays the six relevant cell means.

Black students in the unbuffered criticism and the criticism plus high standards only conditions also showed marginally lower task motivation ($M = -.19$) than did White students in the same conditions ($M = .09$), $t(140) = -1.83, p < .07$. (Introducing quality of students' letters [assessed by a coder unaware of students' condition or race] as a covariate raised the relevant t value to $-2.35, p < .02$.) By contrast, Black students in the wise criticism condition were somewhat more motivated at their task ($M = .27$) than were their White peers ($M = -.06$), $t(140) = 1.54$. (Introducing quality of students' letters as a covariate lowered the relevant t value, $t < 1$.) Evidence for the hypothesized Race \times Condition interaction effect is provided by a weighted contrast that compared the race difference in the unbuffered criticism condition and criticism plus high standards only condition with that in the wise criticism condition, $t(140) = 2.30, p < .05$.

Identification With Writing Skills

An ANCOVA was performed on our measure of general identification with writing skills (using premanipulation identification as the covariate). On this measure, gender interacted significantly with race, such that Black males showed lower identification with writing skills than did Black females, whereas the reverse was true for White students, $F(1, 128) = 11.42, p < .001$. Otherwise, there were no significant main or interaction effects involving race or criticism condition (all p s $> .20$). In par-

ticular, Black and White students showed similar levels of identification with writing skills both in the wise criticism condition and in the criticism plus high standards only condition. However, a planned contrast comparing Black students' identification with writing skills in the unbuffered criticism condition ($M = -0.20$) with their identification in the wise criticism condition ($M = 0.01$) yielded a marginally significant t value, $t(128) = -1.77, p < .08$. (One could argue for a one-tailed test, which would yield a p value of .04, because we are replicating an effect obtained in Study 1.) In addition, the level of general identification with writing skills by Black students in the unbuffered criticism condition proved to be marginally lower than that of White students in the same condition ($M = 0$), $t(128) = -1.67, p < .10$.

Racial Differences in

Performance: A Potential Mediator?

On average, the letters of Black students were rated less positively than those produced by White students (again, as assessed by a coder unaware of students' race or condition). Accordingly, we conducted further analyses to determine whether this difference in performance (and perhaps even a resulting difference in harshness of feedback or in anticipated difficulty of achieving a satisfactory revision) could have played a role in our main findings. When we simply control for differences in rated letter quality in an ANCOVA, the Race \times Feedback interactions most relevant to our predictions remain statistically significant, both for task motivation and ratings of bias. Indeed, when the analysis controls for performance, the relevant contrast involving the identification measure for Black students in the unbuffered criticism condition versus Black students in the wise criticism condition becomes stronger. Furthermore, when we perform a regression analysis including letter quality as a continuous variable, we find no trace of a three-way interaction among race, feedback, and performance. Thus, the simple Race \times Condition interactions we report remain constant across levels of letter-writing performance.

Moreover, the racial difference in performance logically could work against our hypothesis rather than for it. That is, the receipt of harsher criticism, or other evidence of weaker performance, could enhance rather than attenuate students' motivation to revise—especially insofar as students who performed less well might feel that they could more easily improve their performance. Indeed, among both Black and White students across all conditions, lower performance did in fact prove to be significantly associated with greater postmanipulation task motivation, $r(145) = .21, p = .01$. Black students' lower performance, therefore, should have led them to increase their motivation more than White students, not

less. In sum, students' level of performance does not qualify our main findings—that is, whereas the invocation of high standards, and accompanying assurance, exerted little if any influence on White students, this intervention significantly raised motivation and reduced perceptions of bias among Black students.

DISCUSSION

The results of Study 2 confirmed and extended those of Study 1. When negative feedback was presented without additional information, Black students again responded less favorably than did White students. When the same feedback was accompanied by an explicit invocation of high standards and an assurance that the student was capable of meeting those standards, Black students again responded as favorably, if not more favorably, than did White students. This effect was seen on ratings of bias and task motivation, and to a lesser extent, in measures assessing identification with writing skills more generally. Furthermore, whereas an invocation of high standards alone reduced attributions of bias among Black students, it was insufficient to raise their task motivation.

GENERAL DISCUSSION

Minority students may wonder whether they are being viewed through the lens of a stereotype rather than judged on their own merits and recognized for their full potential. The challenge facing the mentor is to provide critical feedback in a manner that discourages attributions of bias and that refutes the threatening limitation alleged by the stereotype. Although our studies focus on minority students, this challenge arises in any context where students face group-based doubts about their abilities or "belonging" within a given domain of achievement. Thus, as Spencer et al. (1999) and Steele (1997) argue, stereotype threat also affects women working in math and the physical sciences.

The present studies offer two findings that speak to this challenge. First, we showed that whereas unbuffered criticism neither prompted attributions of bias nor undermined motivation on the part of White students, such criticism had precisely those effects on Black students. Second, we demonstrated that a wise, theory-informed intervention—that is, one that both invoked high standards and assured students of their capacity to reach those standards—eliminated this race difference in response to criticism. In fact, the motivation of Black students provided with criticism in this wise manner improved so dramatically that it slightly surpassed that of their White peers. Study 1 further showed this wise feedback to be more effective than the common tactic of buffering criticism with nonspecific praise; Study 2 dem-

onstrated that although the invocation of high standards in the absence of the assurance could deflect attributions of racial bias, it could not fully raise Black students' motivation.

Before discussing these results further, one specific finding from Study 2 merits further comment. Whereas the invocation of high standards alone did not sustain Black students' task-specific motivation, it did suffice to maintain their identification with the general domain of writing skills. It is possible that our highly selected students may have been more secure about their general academic abilities than about their capacity to excel at the specific novel task that our studies introduced. It also is possible that the invocation of high standards, even in the absence of personal assurance, created a context in which the negative feedback had fewer esteem-threatening consequences (see Bruner, 1996; Lepper, 1988; Lepper et al., 1990, 1993; Steele, 1992, 1997). The explicit reference to high standards allowed students to treat the criticism not as diagnostic of an underlying lack of ability on their part, or even of a failure in their particular performance, but rather as a reflection of the exacting standards of their reviewer.

Although such an external attribution would protect students' self-esteem, and thereby forestall any need for defensive disidentification from the achievement domain, we believe it left intact their apprehension concerning future performance. That is, Black students in the high standards only condition still had reason to worry about their prospects for future success at the task and, more specifically, their capacity to meet the reviewer's standards. They still faced the discouraging possibility that further effort might meet not with success but with failure and that they might thereby confirm a negative stereotype—a possibility that might even have become more salient to them when they received criticism regarding their essay. In short, Black students in the high standards only condition may have been free to maintain their identification with the general domain of achievement but discouraged from risking further investment in the specific task at hand.

Consideration of Underlying Processes

Our wise intervention was designed to assure Black students that they would not be judged stereotypically. Black students, we contend, felt that they could trust their critic's motives and that they could safely invest their effort, and even their identity, in the task before them. The invocation of high standards ruled out racial bias as a potential motivation behind the feedback, and the assurance reduced the threat of confirming a group-based limitation in ability. Support for this conceptual

analysis was provided primarily by the specific Race \times Treatment interactions we obtained. Our statistical analyses allowed us to test predictions about the conditions under which Black students would report greater bias than would White students and the conditions under which they would feel less motivation. Our analyses also largely confirmed our predictions about the specific measures and specific conditions that would reveal differences in the response between Black students and White students. In short, we believe that the interaction effects reported in Studies 1 and 2 present persuasive, if not compelling, evidence that wise feedback produced benefits for Black students by attenuating the threat of stigmatization.

Beyond our present conceptual and statistical analyses, there are also two social-psychological theories that can be called on to support the wisdom of assuring students of their capacity to reach a higher performance standard. Bandura's (1977, 1986) self-efficacy theory would point to the obvious benefits of enhancing students' confidence in their capacities. Similar advice would be offered by theorists who have examined the motivational consequences of learned helplessness and of other counterproductive attribution strategies (e.g., Diener & Dweck, 1978; Dweck, 1986; Dweck & Leggett, 1988; Weiner, 1972, 1974). That is, students manifesting learned helplessness could be expected to attribute critical feedback to uncontrollable external factors such as task difficulty or bias on the part of the evaluator, or even worse, to uncontrollable internal factors including their own lack of ability. Affirming students' ability to meet a higher standard, through further effort and incorporation of the changes suggested for improvement, would help to counteract such helplessness.

Although these theories support the wisdom of our wise intervention, they also raise the possibility that a condition that provided students with a boost in task-related efficacy but that made no mention of an elevation in standards might be sufficient to produce the benefits that we have demonstrated and discussed in the present research. At first consideration, an "assurance only" condition (i.e., where the evaluator assures students of their capacity without invoking high standards) might seem timely. On further consideration, however, it becomes clear that although such a condition would bear on theoretical and practical concerns, it would do little to clarify the mediating role of self-efficacy. Assuring Black students that they can do better without explicitly or implicitly conveying the message that doing better means meeting a higher standard would be rather empty and, hence, would constitute a weaker efficacy induction. Indeed, if the student interpreted the assurance as a message that hard work could raise the level of their

performance from utter deficiency to mere adequacy, it might even be threatening and counterproductive. Researchers who want to focus on the role of self-efficacy, we believe, would be better served by testing the effects of a condition in which, prior to the presentation of any criticism, students simply receive positive feedback on their writing skills—or perhaps fortuitously overhear their evaluator praise their abilities. Although such an efficacy boost might help to relieve stereotype threat, we suspect that it would still leave Black students wondering about the critic's motives. Indeed, the additional boost in self-confidence might lead them to find racial bias a more plausible attribution for the feedback—especially if their high confidence is accompanied by the belief that their performance “in absolute terms” has been quite satisfactory.

Again, self-efficacy and learned helplessness are clearly relevant to our general concerns, but these theories cannot, in their simplest form, explain our overall pattern of findings. An adequate theoretical treatment must explain why our wise feedback condition offered benefits unique to Black students. In particular, a satisfactory self-efficacy account of our findings must explain why the motivation of Black students suffered a unique drop in the unbuffered criticism condition and why it was uniquely enhanced in the wise criticism condition. Only by recognizing the significance of race, and the threat of stigmatization, can any theory reasonably address our findings.

Clearly, a self-efficacy account would be reasonable if Black students began the experimental task with lower self-efficacy than White students and thus were in greater need of a boost in efficacy. Our data, however, suggest that this was not the case. Prior to the presentation of the critical feedback, Black students and White students did not differ on any motivational premeasure, including an item that specifically asked students whether they thought they could improve the letter they wrote if given the opportunity to revise it. Moreover, recent investigations (including one national, longitudinal study, and one comprehensive review of the relevant research) find no evidence of preexisting racial differences in academic-related self-efficacy, expectancies for success, or negative attributional styles (see Graham, 1994; Tashakkori, 1993; Tashakkori & Thompson, 1991). Black students, accordingly, should not have benefited uniquely from any manipulation that merely offered a general boost in self-efficacy.

In one important respect, however, conceptual analyses involving stereotype threat, attributional style, and self-efficacy converge. Each suggests that although minority students may not generally be prone to undue doubts about their academic efficacy, or to harmful attributions concerning the causes of academic failure, they

may find such doubts more readily triggered when they are confronted with criticism or when they experience performance frustration. In a sense, Black students' self-efficacy may be reasonably high but nevertheless fragile—at least in domains in which the stereotype poses a relevant explanation for negative feedback or frustration (see, in particular, Study 3 of Steele & Aronson, 1995). We hope that our present analysis invites theorists to consider factors such as stereotype threat and their role in mediating the maintenance of efficacy beliefs over time and across situations. An examination of such factors is crucial to a complete account of individual and group differences in motivation and performance.

In considering other theoretical perspectives, we also believe that part of the effectiveness of wise feedback may indeed lie in the message it conveys about the malleable nature of ability—the message that greater effort will yield performance that surpasses the capacities demonstrated to date (Dweck et al., 1995; Dweck & Leggett, 1988; see also Butler, 1987; Earley, 1986; Koestner, Zuckerman, & Koestner, 1987, 1989; Sansone et al., 1989). Again, however, the fact that we obtained Race \times Treatment interactions rather than main effects of treatments must be addressed. The malleability message, we argue, should be particularly important for students who are targets of ability-stigmatizing stereotypes, because these stereotypes imply that ability (or lack of ability) is a fixed group attribute rather than a malleable aspect of the self. Indeed, as we noted in introducing our studies, the guiding philosophy of many of the most successful programs aimed at minority youth is an emphasis on the malleable nature of academic ability—the message that “intelligence can be taught” (Whimbey, 1975).

Theoretical and Practical Implications

The findings documented here point out the obstacles in the classroom posed by race. The mentor trying to convey feedback to the Black student may face a situation in which there is a possibility of mistrust—a mistrust that arises not from the disposition of mentor or student but from race and the meaning imputed to it in our society. To use the terminology of J. J. Gibson's (1979) ecological approach to perception, race constitutes an “affordance” of mistrust—where an affordance is a situational or contextual factor that invites a particular perception of the object of judgment (see also Baron & Bou-dreau, 1987). The challenge to the wise mentor, accordingly, is to establish a learning environment that “affords” more constructive perceptions.

The need to combat the effects of racial stigmatization does not oblige the educator to withhold critical feedback, to lavish praise, or to otherwise lower perfor-

mance standards in the hope of sustaining student motivation. Rather than altering the content of instruction, the educator might instead consider modifying the context in which such instruction occurs. In the case of our highly selected Black students, motivation was sustained not by diluting the one-and-a-half pages of criticism offered or by softening its tone. What proved effective was providing the relevant criticism in a context in which its critical nature could be readily attributed to the existence of high and consistent standards and a belief in the student's capacity to reach them.

We have argued that our wise intervention served to buffer students from the sting of negative feedback. But an alternative possibility exists that we think merits further investigation. We suspect that the effectiveness of the two-faceted wise intervention depends on the provision of rigorous feedback. Had the feedback been cursory rather than critical, students might have doubted the sincerity of the reviewer's self-proclaimed high standards. Indeed, the additional assurance might have seemed condescending if it had accompanied milder feedback (see Lepper et al., 1990, 1993). The critical nature of the feedback used in our studies, we contend, helped both to reinforce the message that high standards were being applied and to justify the provision of the assurance. What is more, it is likely that the rigor of the feedback also communicated the critic's interest in helping the student to reach the higher standard. Many students remarked in the debriefing session that they had been impressed by the rigor of the criticism and that seldom in their undergraduate careers had a teacher or professor taken their efforts so seriously. A valuable question for future research, accordingly, is whether our wise intervention requires the additional evidence of attention and personal concern implicit in such feedback. Effective educators and academic programs convey an unflinching faith in their minority students' potential. But at the same time they do not hesitate to call attention to the gap between students' current performance and the level that they could achieve with unstinting effort (see Vygotsky, 1978).

In sustained relationships with students, the wise mentor, of course, does not simply speak of high expectations and a faith in students' potential. He or she also buttresses this message through expenditures of time and effort, by giving detailed attention to the student's performance, and by providing an empowering pattern of feedback over time (e.g., Dweck et al., 1978; see also Aronson & Linder, 1965; Jones, 1963). Outside such a relationship (or in trying to initiate one), the teacher providing critical feedback must find other means to deflect attributions of racial bias and to sustain motivation and academic identification. It was in the absence of such a relationship that our wise (but relatively minor)

intervention shielded Black students from the consequences of negative racial stereotypes.

In a real sense, the challenge for the mentor of minority students, whether in the day-to-day interactions of an enduring relationship, or in single encounters about a specific academic product, is to make explicit the message that is apt to be implicit for at least the more privileged of nonminority students. This message of belonging, and of still untapped intellectual potential, is vital. As *Los Angeles Times* columnist Frank del Olmo (cited in Mathews, 1988) observed, the less effective educators expect little from their Black and Latino students and ultimately see their expectations confirmed in poor grades and high dropout rates. By contrast, the successful teachers of minority students "share one thing in common . . . they truly believe that their students can succeed in school, and they act on that conviction" (p. 289).

NOTES

1. An additional "bland praise" condition, in which students returning to the second session received a few phrases of nonspecific praise, was later added to this basic 2×3 factorial design. It was included mainly to demonstrate that it was criticism, and not feedback in general, to which Black students would respond differently from White students. Our results confirmed this expectation and also demonstrated that bland praise alone generally proved less effective than wise criticism in motivating students (although the nature of our dependent variables made relevant between-condition comparisons awkward and problematic). No further discussion of this condition will be offered in this necessarily brief report of our research.

2. Pairs comprised either 2 White participants or a White and a Black participant but never 2 Black participants. Because of the relatively small number of Black students on campus, we worried that the presence of 2 Black students might evoke suspicions regarding the study's concern with race. Thorough postexperimental debriefing revealed no student (with the exception of the 1 Black student, noted earlier, whose data were discarded) who voiced any suspicion regarding the significance of race in our study.

3. With one exception, the individual items constituting the task motivation composite were highly correlated (ranging from $r = .40$ to $r = .65$), a result reflected in the principal component analyses reported in our Results sections. The item "How important is it to you to do a good job on your letter?" was only weakly correlated with the item "How much do you think you could improve your letter given an opportunity to revise it?" ($r = .16$). Accordingly, this item was dropped in Study 2. We should note, however, that excluding this item in Study 1 would not change the statistical significance of any effects involving the relevant composite measure.

4. Several potential covariates also were measured. These included students' self-reported perceptions of the quality of their letter, students' self-esteem (assessed by Rosenberg's, 1965, self-report instrument), and students' self-reported verbal and math Scholastic Aptitude Test (SAT) scores. None of the assessed covariates proved a consistent and significant predictor or possible mediator of the effects to be reported, and accordingly, none receives further consideration in this article.

5. One outlier (a White student in the unbuffered criticism condition) provided a bias rating more than 3 standard deviations away from his group mean. We opted to omit his data for analysis of the bias item. Doing otherwise—for example, "Windsorizing" his response (by reducing it to equal the next most extreme score in his cell) or even including his response without alteration—would have lowered the significance of our omnibus analysis but left intact the interaction effect most relevant to our hypothesis. Resorting entirely to nonparametric analyses similarly would have lowered significance levels but still con-

firmed the relevant interaction effect involving race and feedback condition.

REFERENCES

- Aronson, E., & Linder, D. (1965). Gain and loss of esteem as determinants of interpersonal attractiveness. *Journal of Experimental Social Psychology, 1*, 156-171.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review, 84*, 191-215.
- Bandura, A. (1986). *Social foundations of action: A social-cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Banks, W. C., Stitt, K. R., Curtis, H. A., & McQuarter, G. (1977). Perceived objectivity and the effects of evaluative reinforcement upon compliance and self-evaluation in Blacks. *Journal of Experimental Social Psychology, 13*, 452-463.
- Barker, G. P., & Graham, S. (1987). Developmental study of praise and blame as attributional cues. *Journal of Educational Psychology, 79*, 62-66.
- Baron, R. M., & Boudreau, L. A. (1987). An ecological perspective on integrating personality and social psychology. *Journal of Personality and Social Psychology, 53*, 1222-1228.
- Bruner, J. S. (1962). *On knowing: Essays for the left hand*. Cambridge, MA: Harvard University Press.
- Bruner, J. (1996). *The culture of education*. London: Harvard University Press.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivation perceptions, interest, and performance. *Journal of Educational Psychology, 79*, 474-482.
- Comer, J. P. (1980). *School power: Implications of an intervention project*. New York: Free Press.
- Comer, J. P. (1988). Educating poor minority children. *Scientific America, 259*(5), 42-48.
- Cose, E. (1997). *Color-blind: Seeing beyond race in a race-obsessed world*. New York: HarperCollins.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96*, 608-630.
- Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 60*, 218-228.
- Daloz, L. A. (1986). *Effective teaching and mentoring*. San Francisco: Jossey-Bass.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- Diener, C. I., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy and achievement cognitions following failure. *Journal of Personality and Social Psychology, 36*, 451-462.
- Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.
- Dweck, C., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry, 6*, 267-285.
- Dweck, C., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom. *Developmental Psychology, 14*, 268-276.
- Dweck, C., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.
- Earley, P. C. (1986). An examination of the mechanisms underlying the relation of feedback to performance. *Academy of Management Journal, 33*, 87-105.
- Ericsson, K. A. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.
- Fernandez, C., Espinosa, R., & Dornbusch, S. (1975). *Factors perpetuating the low academic status of Chicano high school students* (Memorandum No. 138). Stanford, CA: Stanford University, Center for Research and Development in Teaching.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Goffman, E. (1963). *Stigma: Notes on the management of a spoiled identity*. Englewood Cliffs, NJ: Prentice Hall.
- Graham, S. (1994). Motivation in African Americans. *Review of Educational Research, 64*, 55-117.
- Harber, K. (1998). Feedback to minorities: Evidence of a positive bias. *Journal of Personality and Social Psychology, 74*, 622-628.
- Horn, T. S. (1985). Coaches' feedback and changes in children's perceptions of their physical competence. *Journal of Educational Psychology, 77*, 174-186.
- Jones, E. E. (1963). *Ingratiation: A social psychological analysis*. New York: Appleton Century Crofts.
- Kleck, R. (1966). Physical stigma and nonverbal cues emitted in face-to-face interaction. *Human Relations, 21*, 19-28.
- Kleinfeld, J. (1975). Effective teachers of Eskimo and Indian students. *School Review, 83*, 301-344.
- Koestner, R., Zuckerman, M., & Koestner, J. (1987). Praise, involvement and intrinsic motivation. *Journal of Personality and Social Psychology, 53*, 383-390.
- Koestner, R., Zuckerman, M., & Koestner, J. (1989). Attributional focus of praise and children's intrinsic motivation: The moderating role of gender. *Personality and Social Psychology Bulletin, 15*, 61-72.
- Lepper, M. R. (1988). Motivational considerations in the study of instruction. *Cognition and Instruction, 5*, 289-309.
- Lepper, M. R., Aspinwall, L. G., & Mumme, D. L. (1990). Self-perception and social-perception processes in tutoring: Subtle social control strategies of expert tutors. In J. M. Olson & M. P. Zanna (Eds.), *Self-inference processes: The Ontario symposium* (Vol. 6, pp. 217-237). Hillsdale, NJ: Lawrence Erlbaum.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75-105). Hillsdale, NJ: Lawrence Erlbaum.
- Massey, G. C., Scott, M. V., & Dornbusch, S. M. (1975, November). Racism without racists: Institutional racism in urban school. *Black Scholar*, pp. 10-19.
- Mathews, J. (1988). *Escalante: The best teacher in America*. New York: Henry Holt and Company.
- Meyer, W. U., Bachmann, M., Biermann, U., Hempelmann, M., Ploger, F. O., & Spiller, H. (1979). The informational value of evaluative behavior: Influences of praise and blame on perceptions of ability. *Journal of Educational Psychology, 71*, 259-268.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*, 33-52.
- Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of own attainment, and the understanding that difficult tasks require more ability. *Child Development, 49*, 800-814.
- Prentice, D. A., Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus. *Journal of Personality and Social Psychology, 64*, 243-256.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880-892.
- Sansone, C., Sachau, D. A., & Weir, C. (1989). Effects of instruction on intrinsic interest: The importance of context. *Journal of Personality and Social Psychology, 57*, 819-829.
- Spencer, S., Steele, C. M., & Quinn, D. (1999). Under suspicion of inability: Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.
- Steele, C. M. (1992, April). Race and the schooling of black Americans. *Atlantic Monthly*, 68-78.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance of women and African-Americans. *American Psychologist, 52*, 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797-811.
- Stoutemeyer, K., & Steele, C. M. (1997). *The effects of stereotype threat on women's self-esteem, identification with math, and math performance*. Unpublished manuscript.

- Tashakkori, A. (1993). Race, gender and pre-adolescent self-structure: A test of construct-specificity hypothesis. *Personality and Individual Differences, 14*, 591-598.
- Tashakkori, A., & Thompson, V. D. (1991). Race differences in self-perception and locus of control during adolescence and early adulthood: Methodological implications. *Genetic, Social, & General Psychology Monographs, 117*, 113-152.
- Treisman, P. U. (1985). *A study of the mathematics performance of Black students at the University of California, Berkeley*. Unpublished manuscript.
- Vygotsky, L. V. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Weiner, B. (1972). *Theories of motivation: From mechanism to cognition*. Chicago: Markham.
- Weiner, B. (1974). *Achievement motivation and attribution theory*. Morristown, NJ: General Learning Press.
- Weisinger, H. (1990). *The critical edge: How to criticize up and down your organization and make it pay off*. New York: Perennial Library.
- Whimbey, A. (1975). *Intelligence can be taught*. New York: E. P. Dutton.

Received June 30, 1997

Revision accepted June 30, 1998