

11-18-2016

The Perceived Validity of Stereotype Threat as an Explanation for Underperformance

Gabriel Camacho

University of Connecticut - Storrs, gabriel.camacho@uconn.edu

Recommended Citation

Camacho, Gabriel, "The Perceived Validity of Stereotype Threat as an Explanation for Underperformance" (2016). *Master's Theses*. 1034.
http://digitalcommons.uconn.edu/gs_theses/1034

This work is brought to you for free and open access by the University of Connecticut Graduate School at DigitalCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of DigitalCommons@UConn. For more information, please contact digitalcommons@uconn.edu.

Running head: THE PERCEIVED VALIDITY OF STEREOTYPE THREAT

The Perceived Validity of Stereotype Threat as an Explanation for Underperformance

Gabriel Camacho

B.A., University of Virginia, 2012

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of
Science at the University of Connecticut

2016

APPROVAL PAGE

Master of Science Thesis

The Perceived Validity of Stereotype Threat as an Explanation for Underperformance

Presented by

Gabriel Camacho, B.A.

Major Advisor _____
Diane M. Quinn

Associate Advisor _____
Felicia Pratto

Associate Advisor _____
Stephanie Milan

University of Connecticut

2016

Abstract

Research demonstrates that members of stigmatized groups may underperform when stereotype threat is induced. No research, however, has examined whether attributing underperformance to stereotype threat is perceived as a likely or valid explanation. In two experiments White students were randomly assigned to review the test performance of a fellow student described as a racial/ethnic minority (stigmatized group) or White (non-stigmatized group). Experiment 1 revealed that Hispanic and Black students were significantly more likely to have their underperformance attributed to dispositional and stereotypic causes than a White student. Experiment 2 revealed that White students perceived attributing underperformance to stereotype threat as a less valid explanation of underperformance compared to test anxiety. Moreover, all situational attributions for underperformance were perceived as less valid when provided by a Hispanic rather than White student. These findings suggest that majority students who are less vulnerable to stereotype threat do not detect it in others and are generally incredulous of its adverse effects.

The Perceived Validity of Stereotype Threat as an Explanation for Underperformance

Research demonstrates that members of stigmatized groups are vulnerable to experiencing stereotype threat—the concern of confirming a negative stereotype about one’s group (Steele & Aronson, 1995). This concern can become so overwhelming that it can hinder test performance in the lab (Steele & Aronson, 1995), classroom (Good, Aronson & Harder, 2008), and on state-wide standardized tests (Good, Aronson, & Inzlicht, 2003). Knowing the adverse effects stereotype threat can have on the performance of members of stigmatized groups, it is surprising that no research, to my knowledge, has examined whether attributing underperformance to stereotype threat is perceived as a valid explanation by others. This is even more surprising considering that stereotype-threatened students can be consciously aware of the concern of confirming a negative stereotype (Marx & Goff, 2005), are able to report that this concern contributed to anxiety while completing a test (Johns, Schmader, & Martens, 2005), and have been shown to believe that anxiety can reduce test performance (Boucher, Rydell, & Murphy, 2015). If stereotype-threatened students are able to report that they underperformed on a test due to the anxiety induced by the concern of being perceived as stereotypic, it is important to examine the possible social consequences and perceived validity of attributing underperformance to stereotype threat. This is the purpose of the current research.

Recent research on forecasting the experience of stereotype-threatened women has shown that students not experiencing threat (nonstereotype-threatened students) do have some insight into the experiences of stereotype-threatened students (Boucher et al., 2015). Specifically, both men and women forecasters acknowledged, but overestimated, the increased anxiety stereotype-threatened women experienced while completing a stereotype threat inducing test. Although forecasters acknowledged that stereotype-threatened women would experience increased anxiety,

they did not predict that this anxiety would influence their test performance because of the belief that stereotype-threatened women would overcome the anxiety. If nonstereotype-threatened students are able to acknowledge the experience of stereotype threat but are incapable of predicting its adverse effects, then how would attributions of underperformance due to stereotype threat be perceived? Nothing in the stereotype threat literature directly addresses this question; however, research on the ultimate attribution error and the social cost of attributing negative outcomes to discrimination may provide some valuable insight.

The ultimate attribution error postulates that outgroup members are more likely than ingroup members to have a negative behavior attributed to dispositional causes (Hewstone, 1990; Pettigrew, 1979). These dispositional attributions are influenced by the negative stereotypes associated with an outgroup members social group (Froehlich, Martiny, Deaux, & Mok, 2015). Stereotype-threatened students are often underrepresented in the academic domain in which they are negatively stereotyped. As a result of their outgroup status and the salience and relevance of the negative stereotypes associated with their social group, it is likely that their underperformance will be attributed to dispositional and stereotypic causes—that is, causes that invoke negative stereotypes and are perceived as dispositional to their social group. In line with the ultimate attribution error, the underperformance of negatively stereotyped outgroup members, but not nonstereotyped ingroup members, should be attributed to dispositional and stereotypic causes but not external causes such as stereotype threat.

The ultimate attribution error suggests that stereotype-threatened students will have their underperformance attributed to dispositional causes rather than external causes. Perhaps more informative on the reaction others may have to attributing test performance to stereotype threat is research on reactions to attributing negative outcomes to discrimination. Specifically, research

has found that Black students are more likely to be perceived as complainers when they attribute underperformance to discrimination rather than another external cause (e.g., difficulty of the test) or an internal cause (e.g., lack of ability; Kaiser & Miller, 2001). This remained true even when White students were led to believe that it was certain that the Black student was discriminated against. It has been suggested that attributing negative outcomes to discrimination is disapproved of by members of non-stigmatized groups because, since they are less likely to experience discrimination, they are more likely to underestimate its occurrence and adverse effect (Kaiser & Miller, 2001). Moreover, claiming discrimination may be perceived as excuse making (Kaiser & Miller, 2003).

Attributing underperformance to stereotype threat may similarly be disapproved of by majority students who are less vulnerable to stereotype threat. Although stereotype threat is a situational predicament that anyone can experience (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995), members of stigmatized and underrepresented groups are likely to experience it much more frequently in their day to day lives. Moreover, groups that tend to be the most vulnerable to stereotype threat also tend to be the most consciously aware of the experience of stereotype threat (Spencer, Logel, & Davies, 2016). As a result, majority students who are less vulnerable to stereotype threat may underestimate the negative affect it can have on the performance of stereotype-threatened students because of their lack of personal experience or awareness of experiencing stereotype threat. If majority students are unable to relate to the experience of stereotype-threatened students, they may perceive it as excuse making. Even if stereotype threat is acknowledged by majority students, they may, as Boucher et al. (2015) demonstrated, expect stereotype-threatened students to simply overcome this concern.

Results from these studies suggest that majority students not experiencing stereotype threat will likely attribute the underperformance of stigmatized and underrepresented students to dispositional and stereotypic causes and perceive stereotype threat as an unlikely and invalid explanation for underperformance. The present work examines this possibility. Specifically, I examined how White students—who are not negatively stereotyped in academia and are less vulnerable to stereotype threat—perceived the underperformance and explanations for the underperformance of fellow White students or racial/ethnic minority students—who are negatively stereotyped in academia and are vulnerable to stereotype threat.

Across two studies, participants were randomly assigned to review the test performance of a fellow student described as a racial/ethnic minority or White. Importantly, the tests were completed in a testing environment that has been demonstrated to induce stereotype threat and attenuate the test performance of Hispanic and Black students (stereotype-threatened students) but not that of White students (nonstereotype-threatened students). In Experiment 1, participants provided attributions for the underperformance of a fellow student described as either Hispanic, Black, or White. In Experiment 2, participants rated the perceived validity of one of two explanations given for underperformance—one of them being stereotype threat—ostensibly provided by a fellow student described as either Hispanic or White. Together, the present studies explore the importance of considering how stereotype-threatened students are perceived when they underperform and when they attribute their underperformance to stereotype threat.

Experiment 1

The goal of Experiment 1 was to examine the attributions nonstereotype-threatened students provide for the underperformance of stereotype-threatened students. Research on the ultimate attribution error suggest that the negative behavior (e.g., underperformance) of

negatively stereotyped outgroup members (e.g., Hispanic and Black students) is more likely to be attributed to dispositional causes compared to the negative behavior of an ingroup member (White student). In line with the ultimate attribution error, I hypothesized that White students would be significantly more likely to attribute the underperformance of a fellow student described as Hispanic and Black to dispositional and stereotypic causes compared to a fellow student described as White.

I also examined, on an exploratory basis, whether there were social consequences to underperforming on a stereotype-relevant task. Specifically, I examined whether students that underperform in a domain they are negatively stereotyped in are perceived as less competent and receive differential and substandard feedback compared to students that underperform on the same task but are not negatively stereotyped in the domain. Research has shown that Hispanics and Blacks are generally perceived as less competent compared to their White counterparts (Fiske, Cuddy, Glick, & Xu, 2002). Moreover, research on the self-fulfilling prophecy has long demonstrated that people's behavior towards others is often influenced by their perception of them (Rosenthal, 1994; Rosenthal & Jacobson, 1992). If the underperformance of stereotype-threatened students is more likely to be attributed to dispositional and stereotypic causes, as I hypothesized, then it is reasonable that these students will be perceived as less competent and the feedback these students are provided will be influenced by those attributions. Thus, I predict that fellow students described as Hispanic or Black (stereotype-threatened students) will be perceived as less competent, receive less helpful feedback, and receive less additional feedback compared to fellow students described as White (nonstereotype-threatened students).

Methods

Participants and Design

Two hundred and fifty students attending a large public university in the northeast participated in exchange for course credit. Two participants were excluded from data analysis. One participant completed the experiment twice, so the second of this participant's two data entries was excluded. A second participant was excluded because of reported skepticism of the experimental cover story. Removing these two students resulted in a final sample of 248 participants (56.5% female; $M_{\text{age}} = 18.76$, $SD = 1.04$). Only students who self-identified as White on a prescreen survey were allowed to participate. Participants were randomly assigned to one of three conditions: Hispanic student condition, Black student condition, or White student condition.

Procedure and Measures

Participants were greeted and consented by one of three White experimenters. Participants either provided consent alone or with up to four other participants. Once consent was provided, they were taken into separate rooms in order to complete the experiment on a computer in private. Participants then read the following cover story:

We are interested in learning whether students can provide honest and critical feedback to fellow students on academic tasks when not in each other's presence. Although research has shown that peer feedback is beneficial for students, with the recent influx of online schooling we are interested in examining whether peer feedback is just as effective when provided online by students who have never met.

Today you will be completing a verbal reasoning exam and then reviewing and providing feedback on the exam performance of a fellow university student who previously completed the exam. Your exam performance will be evaluated by another student in the near future. The feedback you provide will be emailed to the student and this is also how

you will receive your feedback.

The only information that will be shared with students is basic demographics, test scores, and the peer feedback. No identifying information (i.e., name) will be shared with fellow students. Since we are interested in learning whether peer feedback benefits students, we ask that you please be honest and critical with the feedback you provide.

In reality, the student whose test performance participants' evaluated was fabricated. Participants were randomly assigned to either review the test performance of a fellow student described as Hispanic, Black, or White. The cover story was designed to provide participants a convincing explanation for providing their honest feedback and to reduce any possible social desirability concerns. If no identifying information was shared with the fellow student, then, I theorized, students would be more willing to be honest and critical in their feedback.

Race/ethnicity manipulation. Participants were then prompted to identify their race/ethnicity, gender, and academic year. They were informed that their responses to these three demographic questions would be shared with the student reviewing their test performance and, likewise, that they would receive the fellow student's responses to the same three demographic questions. The fellow student was described as either Hispanic, Black, or White (race/ethnicity manipulation). The gender and academic year of the fellow student always matched the participant's gender and academic year.

Stereotype threat manipulation. Next, participants were informed that they would be given 10 minutes to complete a 12-item verbal reasoning test. Test questions were taken from the Graduate Record Examination (GRE). The test was described as being designed to examine people's linguistic abilities and intelligence. Moreover, it was described as indicative of one's verbal reasoning abilities and capable of predicting academic achievement. Research has

demonstrated that prompting students to identify their race/ethnicity prior to completing a test described as diagnostic of one's ability and intelligence can induce stereotype threat (Steele & Aronson, 1995). Although previous research has demonstrated that self-identified White students do not typically experience stereotype threat in this testing environment, it is a situation in which Hispanic and Black students, two of the three fictional fellow students, would potentially experience stereotype threat (Gonzales, Blanton, & Williams, 2002; Steele & Aronson, 1995).

Once participants completed the test they were provided with a fictitious percentile rank purportedly indicating the percentage of examinees who took the test and received a lower score. All participants were informed that they received an 85% ranking on the test and, conversely, the fellow student received a 45% ranking on the same test. All participants were led to believe that they received a higher percentile ranking than their fellow student. This was intended to provide the appearance that the fellow student underperformed on the test.

Manipulation check. After participants viewed their and the fellow student's testing percentile, they were prompted to indicate how difficult they perceived the test to be for themselves and for the fellow student on a 7-point Likert scale ranging from 1(*very difficult*) to 7(*very easy*). Participants were also prompted to indicate their perceived performance relative to the fellow student and their perception of the fellow student's performance relative to them on a 7-point Likert scale ranging from 1(*Much Worse*) to 7(*Much Better*). These questions were intended as a check to make sure that students realized that the other student had underperformed.

Attributions for test performance. Participants were then prompted to respond to two open-ended questions requesting that they explain why they performed the way they did on the test (seemingly well), and why the fellow student performed the way he or she did on the test

(noticeably underperformed). Two research assistants coded participants' responses for varying attributions for the fellow student's test performance¹

Coding categories were created for participants' attributions for the fellow student's test performance by looking for clear themes or categories in the attributions provided. Five unique coding categories were found: *academic interests*, *language*, *education*, *test difficulty*, and *lack of effort*. A response was coded as attributing performance to (1) *academic interests* if it explicitly stated that the fellow student's underperformance was due to being more interested in academic domains other than the one he or she was currently tested on (e.g., "Maybe she is a STEM major and is better at math"). A response was coded as attributing performance to (2) *language* if it explicitly stated that the fellow student's underperformance was due to speaking English as a second language (e.g., "English might not be their first language"). A response was coded as attributing performance to (3) *education* if it explicitly stated that the fellow student's underperformance was due to having a substandard education (e.g., "Maybe she did not have the same type of education background that I have had"). A response was coded as attributing performance to (4) *test difficulty* if it explicitly stated that the fellow student's underperformance was due to the test being difficult (e.g., "The test was challenging in its own right"). Lastly, a response was coded as attributing performance to (5) *lack of effort* if it explicitly stated that the fellow student's underperformance was due to the fellow student not trying (e.g., "He didn't really care about his results"). All attributions were coded dichotomously (0 = attribution not mentioned or 1 = attribution mentioned). Interrater reliability was moderate to high for all

¹ Since the intended purpose of the experiment was to examine the self-generated attributions for the underperformance of others rather than the attributions for one's own stellar test performance, participants' attributions of their own performance was not coded.

attributions, Cohen's Kappa $\geq .79$. Cases of disagreement were resolved by a third coder. See Table 1 for the percentage of overall attributions falling into each category.

Perceived competence. Next students responded to a modified version of the competence scale from the stereotype content model (Fiske et al, 2002). This scale was used to access participants' perception of the fellow student's competence. The modified competence scale was intended to indicate whether having one's underperformance attributed to dispositional and stereotypic factors also leads to one being perceived as less competent. To ensure that participants' perception of the fellow student's competence was being measured and not the participants' perception of how the fellow student would be perceived in society—as the original scale does—items were changed from asking how society would view the fellow student to asking how the participant viewed the fellow student (e.g., How competent do you think s/he is?). The five competence scale items were rated using a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*extremely*). The internal consistency of the modified competence scale was adequate, $\alpha = .80^2$.

Peer Feedback. Participants were then given two minutes to provide honest, critical, and helpful feedback to the student whose test performance they reviewed. Participants were informed that the fellow student would read the feedback provided. Two research assistants coded participants' responses for varying types of feedback. Coding categories were created for peer feedback by looking for clear themes or categories in the feedback provided. Five unique

² After completing the competence scale, participants responded to three author-developed questions that were intended to access participants' perception of the fellow student in academia. However, the internal consistency of these three questions was low, $\alpha = .63$. As result, this scale will not be discussed for the remainder of Experiment 1.

coding categories were found: *read more*, *expand vocabulary*, *try harder*, *test-taking strategies*, and *no strategies* provided.

Feedback was coded as suggesting to (1) *read more* if it explicitly stated that the fellow student could improve his or her future test performance by reading more (e.g., “Read more books!”). Feedback was coded as suggesting to (2) *expand one’s vocabulary* if it explicitly stated that the fellow student could improve his or her future test performance by learning more vocabulary (e.g., “I think maybe you need a little more practice with vocab”). Feedback was coded as suggesting to (3) *try harder* if it explicitly stated that the fellow student could improve his or her future test performance by asserting more effort on the test (e.g., “Put 100 percent into it if you are going to take your time out to participate in something”). Feedback was coded as providing (4) *test-taking strategies* if it explicitly suggested methods for answering test questions (e.g., “Try putting the word options in the sentence in your head to see if they sound right”). Lastly, feedback was coded as providing (5) *no strategies* if it did not explicitly state any suggestions for improving future test performance (e.g., “Get them next time”). Feedback was coded dichotomously (0= category of feedback not provided or 1=category of feedback provided). Interrater reliability was moderate to high for all feedback, Cohen’s Kappa, $\geq .73$. Cases of disagreement were resolved by a third coder. See Table 2 for the percentage of overall feedback falling into each category.

Lastly, after providing feedback, participants were asked “How likely are you to spend some additional time today providing helpful feedback to the student?”. Participants responded to this question on a 7-point Likert scale ranging from 1 (*Very Unlikely*) to 7 (*Very Likely*). It was expected that because students were only given two minutes to provide feedback, that a substantial percentage of them would be willing to spend more time providing additional

feedback to the fellow student. However, I hypothesized that participants would be less willing to provide additional feedback to fellow students described as Hispanic or Black because additional feedback would be perceived as less helpful for students whom underperformance was attributed to dispositional and stereotypic causes.

Results

Manipulation Check

A one-way repeated-measures ANOVA was calculated comparing participants' perception of the difficulty of the test for themselves and the fellow student. A significant effect was found, $F(1,247) = 43.15, p < .001, \eta^2_p = .15$. Participants perceived the test to be more difficult for the fellow student ($M = 1.95, SD = .89$) than for themselves ($M = 2.44, SD = .98$). A one-way repeated-measures ANOVA comparing participants' perception of their test performance and the fellow student's test performance also revealed an effect, $F(1,247) = 317.91, p < .001, \eta^2_p = .56$. Specifically, participants believed that the fellow student performed worse on the test ($M = 2.48, SD = 1.24$) compared to themselves ($M = 5.46, SD = 1.57$). Thus, the manipulation of test performance and feedback was successful.

Attribution for Underperformance

A chi-square test of independence was conducted to examine the effect race/ethnicity of the fellow student had on participants' attributions for the fellow student's underperformance (Figure 1). Results revealed that underperformance was more likely to be attributed to *language* when the fellow student was described as Hispanic (23.7%) rather than White (5.3%) or Black (3.8%), $\chi^2(2, N=230) = 19.33, p < .001, v = 0.29$. Underperformance was also more likely to be attributed to *education* when the fellow student was described as Black (30.4%) rather than Hispanic (17.1%) or White (14.7%), $\chi^2(2, N=230) = 6.73, p = .04, v = 0.17$. Moreover,

underperformance was more likely to be attributed to *test difficulty* when the fellow student was described as White (28.0%) rather than Hispanic (15.8%) or Black (11.4%), $\chi^2(2, N=230) = 7.57, p = .02, v = 0.18$. Lastly, underperformance was more likely to be attributed to *lack of effort* when the fellow student was described as White (30.7%) rather than Hispanic (18.4%) or Black (15.2%), $\chi^2(2, N=230) = 6.06, p = .05, v = 0.16$. There was no significant main effect of race/ethnicity on attributing underperformance to *academic interests*, $\chi^2(2, N=230) = .57, p = .75, v = .05$.

Perceived Competence

A one-way between participant ANOVA was conducted to examine the effect race/ethnicity of the fellow student had on participants' perception of the fellow student's competence. The main effect of race on perceived competence did not reach statistical significance, $F(2, 236) = 2.32, p = .10, \eta^2_p = .019$. On average, participants perceived the fellow student to be of average competence ($M = 2.95, SD = .53$).

Peer Feedback

A chi-square test of independence was conducted to examine the effect race/ethnicity of the fellow student had on the feedback participants' provided. Results revealed that *test-taking strategies* were more likely to be provided as feedback for fellow students described as White (48.7%) rather than Black (34.1%) or Hispanic (30.3%), $\chi^2(2, N=236) = 6.26, p = .04, v = 0.16$ (Figure 2). No significant main effects were revealed between the race/ethnicity of the fellow student and providing *no strategies* as feedback or providing feedback suggesting to *read more, expand one's vocabulary, or try harder* (all $ps > .05$).

Lastly, a one-way between participant ANOVA was conducted to examine the effect race/ethnicity of the fellow student had on participants' willingness to provide additional

feedback. Results revealed no significant main effect of race/ethnicity on willingness to provide additional feedback, $F(2, 244) = .62, p = .54, \eta^2_p = .005$. On average, participants rated themselves as somewhat unlikely to provide additional feedback to the fellow student ($M = 3.43, SD = 1.70$).

Discussion

Experiment 1 examined White nonstereotype-threatened students' attributions for the underperformance of minority stereotype-threatened students. As hypothesized, the underperformance of stereotype-threatened students was more likely to be attributed to dispositional and stereotypic causes compared to the underperformance of nonstereotype-threatened students. Specifically, fellow students described as Hispanic were more likely to have their underperformance attributed to *language* and fellow students described as Black were more likely to have their underperformance attributed to *education*. Both of these attributions perpetuates the negative stereotypes that Hispanics are poor English speakers (Mastro & Behm-Morawitz, 2005) and that Blacks are undereducated (Devine, 1989).

However, the underperformance of nonstereotype-threatened students was more likely to be attributed to situational causes compared to the underperformance of stereotype-threatened students. Fellow students described as White were more likely than fellow students described as Hispanic or Black to have their underperformance attributed to *test difficulty* and *lack of effort*. These attributions are related to the test and testing environment and simply acknowledges that the test was challenging and fellow student's may not have taken the test seriously. Interestingly, there were no differences in attributing underperformance to academic interests. This may be because, at least in American culture, it is socially acceptable to be bad at math but not in other academic domains (Beilock & Willingham, 2014). This may also explain why there were so few overall attributions to academic interests.

Experiment 1 also explored whether there were social consequences to underperforming on a stereotype-relevant task. Contrary to my hypothesis, stereotype-threatened students were not perceived as less competent compared to nonstereotype-threatened students. This may be because participants did not perceive test performance as a measure of general competence. It is also possible that modifying the scale to inquire about participants' perception about a specific student's competence, rather than ask about societies' perception of varying social groups, engendered social desirability and consequently caused participants to not answer truthfully.

Also as hypothesized, stereotype-threatened students received differential and substandard feedback compared to nonstereotype-threatened students. Specifically, fellow students described as White were more likely to receive *test-taking strategies*—arguably the most helpful form of feedback—compared to fellow students described as Hispanic or Black. Considering that research has shown how our expectations of people influence our behavior toward them (Rosenthal, 1995; Rosenthal & Jacobson, 1992), it seems reasonable that students whose underperformance was attributed to *test difficulty* would be more likely to receive *test-taking strategies* as feedback because it addresses the perceived cause for their underperformance. Under this reasoning, students whose underperformance was attributed to dispositional and stereotypic causes would be less likely to receive *test-taking strategies* as feedback because such strategies would not change or improve the perceived cause for their underperformance.

Lastly, and contrary to my hypothesis, participants were not less willing to provide additional feedback to stereotype-threatened students compared to nonstereotype-threatened students. Although I assumed that most students would be willing to spend more time providing additional feedback, the majority of participants indicated that they were somewhat unlikely to

do so. As a result, the construct that the question assessed was different than what was intended. Specifically, the intended purpose of the scale was to assess whether participants would be less likely to provide additional feedback to stereotype-threatened students; however, since most participants were generally unwilling to provide additional feedback, the scale actually assessed participants' willingness to provide *more* feedback to stereotype-threatened students. In order to examine whether participants are less willing to provide additional feedback to fellow students described as Hispanic or Black, it is first necessary to find enough participants who are willing to provide additional feedback to anyone.

Overall, Experiment 1 revealed that majority students who are less vulnerable to stereotype threat do not detect its adverse effects in stereotype-threatened students but rather, as suggested by the ultimate attribution error, are more likely to attribute underperformance to dispositional and stereotypic causes. Although stereotype-threatened students were not perceived as less competent compared to nonstereotype-threatened students, they were less likely to receive the most helpful form of feedback. Thus, demonstrating that there are social consequences to underperforming on stereotype-relevant tasks.

Experiment 2

The first goal of Experiment 2 was to replicate the effect race/ethnicity of the fellow student had on the attributions provided for the fellow student's underperformance. Since the attributions participants' provided were spontaneous, it was important to examine whether a new cohort of participants would provide similar attributions and whether these attributions remained dispositional and stereotypic for stereotype-threatened students and situational for nonstereotype-threatened students.

The second and main goal of Experiment 2 was to examine how nonstereotype-threatened students perceive stereotype threat as an explanation for underperformance. If majority students are less likely to experience stereotype threat and cannot detect that others are experiencing the threat, as demonstrated in Experiment 1, they may underestimate its occurrence and its adverse effect. In line with this reasoning, I hypothesized that stereotype threat would be perceived as a less valid explanation for underperformance compared to other psychological threats that similarly reduce test performance through anxiety and cognitive impairment—namely, test anxiety.

Thus, in the second study I manipulated both the race/ethnicity of the fellow student (White or Hispanic) and the explanation the fellow student gave for their underperformance (stereotype threat or test anxiety). However, because White students do not contend with any negative stereotypes associated with verbal reasoning ability, the stereotype threat explanation for the Hispanic and White fellow student was intentionally different in order to reflect the disparate stereotypes and concerns both social groups may experience. Hispanic students are negatively stereotyped in academia whereas White students are positively stereotyped. As a result, the Hispanic stereotype threat explanation reflected the concern of possibly confirming a negative stereotype; the White stereotype threat explanation reflected the concern of possibly being unable to confirm a positive stereotype. Although no research has demonstrated that White students experience stereotype threat when they are unable to confirm a positive stereotype about their racial group, research has shown that different social groups experience different forms of stereotype threat (Shapiro, 2011) and positive stereotypes can negatively impact performance (Cheryan & Bodenhausen, 2000). I hypothesized that explanations for underperformance would

be perceived as less valid when it was provided by a fellow students described as Hispanic rather than White.

I again examined, on an exploratory basis, whether there were social consequences to underperforming on a stereotype-relevant task. I examined whether students that underperformed in a domain they are negatively stereotyped in are perceived as having lower verbal ability and receive differential and substandard feedback compared to students who underperformed on the task but are not negatively stereotyped in the domain.

Methods

Participants and Design

A total of 256 students attending a large public university participated in exchange for course credit. Two participants were excluded because they were skeptical of the study's cover story, resulting in a final sample of 254 participants (73.6% female; $M_{age} = 18.99$, $SD = 1.14$). Similar with Experiment 1, only students that identified as White were allowed to participate. Experiment 2 used a 2(race/ethnicity for fellow student: White or Hispanic) x 3(explanation for underperformance: no explanation, test anxiety, or stereotype threat) between participant factorial design.

Procedure and Measures

Participants were greeted and consented by one of two White experimenters. Participants either provided consent alone or with up to four other participants. Once consent was provided, they were taken into separate rooms in order to complete the experiment on a computer in private. Participants then read the same cover story as in Experiment 1. As in Experiment 1, the student whose test performance participants' evaluated was fabricated. Participants were randomly assigned to either review the test performance of a fellow student described as White

or Hispanic and to either receive no explanation for the fellow student's underperformance, test anxiety as the explanation for the fellow student's underperformance, or stereotype threat as the explanation for the fellow student's underperformance.

Race/ethnicity manipulation. Participants then received the same prompt as in Experiment 1 to identify their race/ethnicity, gender, and academic year. They were informed that their responses to these three demographic questions would be shared with the student reviewing their test performance and, likewise, that they would receive the fellow student's responses to the same three demographic questions. The fellow student was described as either White or Hispanic (race manipulation). The gender and academic year of the fellow student always matched the participant's gender and academic year.

Stereotype threat manipulation. Next, participants were informed that they would be given 10 minutes to complete a 12-item verbal reasoning tests. As in Experiment 1, the test was described as diagnostic of ability and intelligence and was designed, along with the pre-demographic questions, to create a stereotype threat inducing testing environment.

Once participants completed the test they were provided with a fictitious percentile rank purportedly indicating the percentage of examinees who took the test and received a lower score. As in Experiment 1, all participants were informed that they received an 85% ranking on the test and, conversely, the fellow student received a 45% ranking on the same test.

Manipulation check. After participants viewed their and the fellow student's testing percentile, they were prompted to indicate how difficult they perceived the test to be for themselves and the fellow student. Participants were also prompted to indicate their perceived performance relative to the fellow student and their perception of the fellow student's performance relative to them.

Attributions for test performance. Participants were then prompted to respond to the same two open-ended questions requesting that they explain why they and the fellow student performed the way they did on the test. As in Experiment 1, two research assistants coded participants' responses for varying attributions for the fellow student's test performance. The same five unique coding categories used in Experiment 1 were used in Experiment 2: *academic interests, language, education, test difficulty, and lack of effort*. All attributions were coded dichotomously (0 = attribution not mentioned or 1 = attribution mentioned). Interrater reliability was moderate to high for all attributions, Cohen's Kappa $\geq .80$. Cases of disagreement were resolved by a third coder. See Table 3 for the percentage of overall attributions falling into each category.

Perceived validity of explanation. After participants provided their attributions for the fellow student's test performance, they read, "Below is the other student's response for why s/he performed the way s/he did. Please read it as we will ask you some questions about their explanation." Participants were randomly assigned to read one of two explanations or not provided an explanation at all (control condition). When test anxiety was provided as the explanation for the fellow student's underperformance, participants read, "I tend to feel anxious when taking tests. I basically become so worried that I'm not going to do well on the test that it makes it difficult for me to focus. I think this is why I didn't do so well on the test." The test anxiety explanation was the same for both the Hispanic and White fellow student.

When stereotype threat was provided as the explanation for the White fellow student's underperformance, participants read, "I tend to feel anxious when taking tests. I basically become so worried that I won't do as well as people expect me to do as a White student that it makes it difficult for me to focus. I think this is why I didn't do so well on the test." When

stereotype threat was provided as the explanation for the Hispanic fellow student's underperformance, participants read, "I tend to feel anxious when taking tests. I basically become so worried that if I don't do well on the test people will think it's because I'm Hispanic and this makes it difficult for me to focus. I think this is why I didn't do so well on the test." Again, the stereotype threat explanation for the Hispanic and White fellow student was intentionally different in order to reflect the disparate stereotypes and concerns both social groups may experience.

Participants who read the fellow student's explanation for underperforming on the test were then asked to respond to the following question, "Do you think the student's explanation for his/her test performance is accurate? In other words, do you think that his/her explanation is really the cause for his/her performance (whether s/he performed well or not)?" Participants responded to this questions on a 5-point Likert scale ranging from 1 (*Definitely not*) to 5 (*Definitely yes*). Participants who were randomly assigned to receive no explanation for underperformance (control condition) did not respond to this question.

Cause of perceived validity. In an attempt to better understand participants perceived validity of the provided explanation, they were asked, "please use the space below to explain your answer." Two research assistants coded participants' responses for various explanations for their ranking of validity. Coding categories were created by looking for clear themes or categories in the explanations provided. Two unique coding categories were found: *unlikely* and *reliable*. An explanation was coded as (1) *unlikely* if the participant explicitly stated that the fellow student's explanation for underperforming could have influenced their test performance but was unlikely to be the actual cause of his or her underperformance (e.g., "It is possible but I would say personally that it might be an excuse."). An explanation was coded as (2) *reliable* if

participants explicitly stated that they could relate to the fellow student's explanation for underperforming (e.g., "I can relate to this student because I often feel anxious when taking tests as well."). The cause of perceived validity was coded dichotomously (0 = specific cause not mentioned or 1 = specific cause mentioned). Interrater reliability was high for both explanations, Cohen's Kappa $\geq .81$. Cases of disagreement were resolved by a third coder. See Table 4 for the percentage of overall attributions falling into each category.

Perceived competence. Next students responded to the same modified version of the competence scale used in Experiment 1. Unlike Experiment 1, however, the internal consistency of the modified competence scale was low, $\alpha = .62$. As result, this scale will not be discussed further³.

Perceived verbal ability. Participants were then asked, "Please rate the student's verbal ability" on a 5-point Likert scale ranging from 1 (*very low verbal ability*) to 5 (*very high verbal ability*).

Peer Feedback. Participants were then given two minutes to provide honest, critical, and helpful feedback to the student whose test performance they reviewed. Participants were informed that the fellow student would read the feedback provided. Two research assistants coded participants' responses for varying types of feedback. As in Experiment 1, coding categories were created for peer feedback by looking for clear themes or categories in the feedback provided. Two unique coding categories were found: *test-preparation strategies* and

³Participants then responded to the same three author-developed questions measuring participant's perception of the fellow student in academia. As in Experiment 1, the internal consistency of these three questions was non-existent, $\alpha = -.03$, and will not be discussed further.

anxiety-reducing strategies. See Table 5 for the percentage of overall attributions falling into each category.

Feedback was coded as providing (1) *test-preparation strategies* if it explicitly provided suggestions for preparing for a future test (e.g., “I think that reading more is the only way you can really improve on tests like these.”). Feedback was coded as providing (2) *anxiety-reducing strategies* if it explicitly provided any suggestions for improving test performance by reducing anxiety (e.g., “I believe that you should just try to take your time on the exam, and breathe. Try not to worry about getting the answer right or wrong and focus on what the question is asking you to do and that will lead to your success.”). Feedback was coded dichotomously (0= category of feedback not provided or 1=category of feedback provided). Interrater reliability was moderate to high for all feedback, Cohen’s Kappa, $\geq .76$. Cases of disagreement were resolved by a third coder. Lastly, after providing feedback, participants were asked to respond to the same question in Experiment 1 asking the likelihood of them providing additional feedback to the fellow student.

Results

Manipulation Check

A one-way repeated-measures ANOVA was calculated comparing participants’ perception of the difficulty of the test for themselves and the fellow student. A significant effect was found, $F(1,253) = 48.86, p < .001, \eta^2_p = .16$. Participants perceived the test to be more difficult for the fellow student ($M = 1.95, SD = .95$) than for themselves ($M = 2.44, SD = .87$). A one-way repeated-measures ANOVA comparing participants’ perception of their test performance and the fellow student’s test performance also revealed an effect, $F(1,253) = 335.79, p < .001, \eta^2_p = .57$. Specifically, participants believed that the fellow student performed worse on the test

($M=2.48$, $SD=1.27$) compared to themselves ($M=5.55$, $SD=1.55$). Thus, the manipulation of test performance and feedback was again successful.

Attribution for Underperformance

A chi-square test of independence was conducted to examine the effect race/ethnicity of the fellow student had on participants' attributions for the fellow student's underperformance⁴ (Figure 3). Replicating Experiment 1, underperformance was more likely to be attributed to *language* when the fellow student was described as Hispanic (13.1%) rather than White (0.9%), $\chi^2(1, N=239) = 13.59$, $p < .001$, $OR=17.51$, $CI_{95\%} = [2.28, 134.31]$. Also replicating Experiment 1, underperformance was more likely to attributed to *test difficulty* when the fellow student was described as White (18.8%) rather than Hispanic (8.2%), $\chi^2(1, N=239) = 5.79$, $p = .02$, $OR=.39$, $CI_{95\%} = [.17, .86]$. Moreover, race/ethnicity had no effect on attributions or *academic interests*, $\chi^2(1, N=239) = .14$, $p = .71$, $OR=1.19$, $CI_{95\%} = [.47, 2.98]$. However, unlike Experiment 1, underperformance was more likely to attributed to *education* when the fellow student was described as Hispanic (27.0%) rather than White (14.5%), $\chi^2(1, N=239) = 5.66$, $p = .02$, $OR=2.18$, $CI_{95\%} = [1.14, 4.18]$ and there was no significant main effect of race/ethnicity on attributions to *lack of effort*, $\chi^2(1, N=239) = 1.70$, $p = .19$, $OR=1.54$, $CI_{95\%} = [.80, 2.94]$.

Perceived Validity of Explanation

A 2 (race/ethnicity of fellow student: Hispanic or White) x 2 (explanation provided: test anxiety or stereotype threat) ANOVA was conducted to examine the effect race/ethnicity and

⁴ Participants provided their attributions for the fellow student's underperformance prior to being randomly assigned to either receive an explanation for underperformance or not. Therefore, the chi-square test of independence only examines main effects of race/ethnicity and not explanation.

explanation had on the perceived validity of the provided explanation (Figure 4). Results revealed a significant main effect of race/ethnicity, $F(1, 165) = 4.18, p = .04, \eta^2_p = .025$. As predicted, participants perceived the fellow student's explanation as less valid if the fellow student was described as Hispanic ($M = 3.28, SD = .94$) rather than White ($M = 3.59, SD = 1.0$). Results also revealed a significant main effect of explanation, $F(1, 165) = 13.98, p < .001, \eta^2_p = .078$. As predicted, participants perceived the stereotype threat explanation ($M = 3.18, SD = 1.06$) as less valid compared to the test anxiety explanation ($M = 3.72, SD = .83$). However, results revealed no significant interaction effect between race/ethnicity and explanation, $F(1, 165) = .94, p = .33, \eta^2_p = .006$.

Cause of Perceived Validity

Race/ethnicity, explanation, and their interaction were entered into logistic regression predicting participants' explanation for their perceived validity ranking (Figure 5 & 6). Results revealed that participants were significantly more likely to state that the fellow student's explanation for underperformance was *unlikely* when the fellow student was described as Hispanic (45.5%) rather than White (30.0%), Wald $\chi^2(1) = 3.94, p = .05, OR = 1.95, CI_{95\%} = [1.01, 3.76]$. Participants were also significantly more likely to state that the fellow student's explanation for underperformance was *unlikely* when stereotype threat (48.2%) was provided as the explanation rather than test anxiety (26.2%), Wald $\chi^2(1) = 8.46, p = .004, OR = 2.66, CI_{95\%} = [1.38, 5.13]$. The interaction between race/ethnicity and explanation was not significant, Wald $\chi^2(1) = .224, p = .64, OR = 1.37, CI_{95\%} = [.37, 5.13]$.

Results also revealed that participants were significantly more likely to state that the fellow student's explanation for underperformance was *relatable* when the fellow student was described as White (28.9%) rather than Hispanic (15.6%), Wald $\chi^2(1) = 4.61, p = .03, OR = .40,$

CI_{95%} = [.17, .92]. Participants were also more likely to state that the fellow student's explanation was *relatable* when test anxiety (28.6%) rather than stereotype threat (16.9%) was the provided explanation for underperformance, Wald $\chi^2(1) = 3.92, p = .05, OR = .43, CI_{95\%} = [.19, .99]$. The interaction between race/ethnicity and explanation was not significant, Wald $\chi^2(1) = 1.22, p = .27, OR = .39, CI_{95\%} = [.07, 2.07]$

Perceived Verbal Ability

A 2 (race/ethnicity of fellow student: Hispanic or White) x 3 (explanation provided: no explanation, test anxiety, or stereotype threat) ANOVA was conducted to examine the effect race/ethnicity of the fellow student and explanation provided had on the perceived verbal ability of the fellow student (Figure 7). Results revealed a marginal main effect of race/ethnicity, $F(1, 248) = 3.31, p = .07, \eta^2_p = .013$. Fellow students described as Hispanic ($M = 2.61, SD = .69$) were perceived as having lower verbal ability compared to fellow students described as White ($M = 2.81, SD = .87$). Results revealed no significant main effect of explanation, $F(2, 248) = 1.29, p = .277, \eta^2_p = .010$. Fellow student's that provided stereotype threat as an explanation for underperformance ($M = 2.73, SD = .70$) was not perceived to have a lower verbal ability compared to fellow students that provided test anxiety ($M = 2.81, SD = .84$) as an explanation for underperformance or provided no explanation for underperformance ($M = 2.59, SD = .82$). Results also revealed no significant interaction effect between the fellow student's race/ethnicity and the explanation provided, $F(2, 248) = .26, p = .77, \eta^2_p = .002$.

Peer Feedback

Race/ethnicity, explanation, and their interaction were entered into logistic regression predicting the feedback participants provided (Figure 8 & 9). Results revealed no significant differences in providing *test-preparation strategies* as feedback for fellow students described as

Hispanic (42.1%) or White (35.2%), Wald $\chi^2(1) = .21, p = .65, OR = 1.14, CI_{95\%} = [.65, 2.01]$.

However, participants were significantly more likely to provide *test-taking strategies* as feedback when no explanation (65.0%) was provided for underperformance rather than stereotype threat (27.8%), Wald $\chi^2(1) = 20.91, p < .001, OR = .20, CI_{95\%} = [.10, .40]$ or test anxiety (23.8%), Wald $\chi^2(1) = 26.37, p < .001, OR = .16, CI_{95\%} = [.08, .33]$. The interaction between race/ethnicity and explanation was not significant, Wald $\chi^2(2) = 1.60, p = .45$.

Results also revealed no significant differences in providing *anxiety-reducing strategies* as feedback for fellow students described as Hispanic (46.3%) or White (50.0%), Wald $\chi^2(1) = .02, p = .90, OR = .94, CI_{95\%} = [.37, 2.42]$. However, participants were significantly more likely to provide *anxiety-reducing strategies* as feedback when stereotype threat (73.4%) or test anxiety (66.7%) was provided as an explanation for underperformance rather than no explanation (3.8%), Wald $\chi^2(1) = 41.75, p < .001, OR = 81.58, CI_{95\%} = [21.46, 310.06]$, Wald $\chi^2(1) = 35.71, p < .001, OR = 53.23, CI_{95\%} = [14.45, 196.06]$. The interaction between race/ethnicity and explanation was not significant, Wald $\chi^2(2) = 3.42, p = .18$.

Lastly, a 2 x 3 ANOVA conducted to examine the effect race/ethnicity and explanation had on participants' willingness to provide additional feedback revealed no significant main effects of race/ethnicity, explanation, or their interaction (all $p > .05$).

Discussion

Replicating the results from Experiment 1, stereotype-threatened students (in this study, Hispanic students) were more likely to have their underperformance attributed to dispositional and stereotypic causes compared to nonstereotype-threatened (White) students. Specifically, the underperformance of fellow students described as Hispanic was more likely to be attributed to *language* and *education* whereas the underperformance of fellow students described as White

was more likely to be attributed to *test difficulty*. Although Hispanic students in Experiment 1 were not more likely to have their underperformance attributed to education compared to White students, attributing the underperformance of Hispanic students to *education* perpetrates the negative stereotype that Hispanics are undereducated (Niemann, et al., 1994). Replicating the effect race/ethnicity had on attributions for underperformance further supports the theory that majority students who are less vulnerable to stereotype threat do not detect its adverse effects in stereotype-threatened students but rather are more likely to attribute their underperformance to dispositional and stereotypic causes.

As hypothesized, the provided explanation for underperformance was perceived as less valid when provided by a fellow student described as Hispanic rather than White and when stereotype threat was the provided explanation for underperformance rather than test anxiety. Examining participants' reasoning for their perceived validity of the provided explanation revealed that White nonstereotype-threatened students perceived the provided explanation as *unlikely* to be true when it was provided by a fellow student described as Hispanic rather than White and when stereotype threat was the provided explanation rather than test anxiety. Conversely, White nonstereotype-threatened students *related* more to the provided explanation when it was provided by a fellow White student and when test anxiety was the provided explanation. Thus, suggesting that the perceived validity of stereotype threat as an explanation for underperformance is influenced, at least in part, by participants' ability to relate to the fellow student and his or her experience.

Experiment 2 also explored whether there were social consequences to underperforming on a stereotype-relevant task. Results revealed that fellow students that provided an explanation for underperformance were not perceived as having lower verbal ability compared to fellow

students that provided no explanation for underperformance. However, Hispanic students were perceived as having marginally lower verbal ability compared to fellow students described as White. This marginal effect may stem from stereotypes that Hispanics are poor English speakers.

Providing an explanation for underperformance, however, did influence the feedback participants provided. Specifically, participants tailored their feedback based on the perceived cause of the fellow student's underperformance. When no explanation was provided for underperformance, participants were more likely to provide *test-preparation strategies* to fellow students. When either stereotype threat or test anxiety was provided as an explanation for underperformance, participants were more likely to provide *anxiety-reducing strategies* to fellow students. Although the race/ethnicity of the fellow student had no effect on the feedback participants provided, these results demonstrate, as in Experiment 1, that participants provide feedback that addresses the perceived cause of underperformance. When an explanation is provided for underperformance—even when participants perceive the explanation as unlikely to be true—participants respond to the provided explanation rather than their attributions.

Overall, Experiment 2 revealed that majority students who are less vulnerable to stereotype threat are generally incredulous of stereotype threat's adverse effects. Skepticism of stereotype threat appears to stem from nonstereotype-threatened students' inability to relate to the experience of stereotype-threatened students. Although providing an explanation for underperformance did not significantly affect the fellow student's perceived verbal ability, it did influence the type of feedback participants provided.

General Discussion

The current research tested whether majority students less vulnerable to stereotype threat detect it in others and whether they perceive it as a valid explanation for underperformance.

Experiment 1 found that the underperformance of minority stereotype-threatened students was more likely to be attributed to dispositional and stereotypic causes compared to the underperformance of White nonstereotype-threatened students. Experiment 2 replicated this finding and also revealed that White nonstereotype-threatened students perceived attributing underperformance to stereotype threat as a less valid explanation of underperformance compared to test anxiety. Moreover, all situational attributions for underperformance were perceived as less valid when provided by a Hispanic rather than White student. These findings suggest that majority students who are less vulnerable to stereotype threat do not detect it in others and are generally incredulous of its adverse effects.

These findings are consistent with research on the ultimate attribution error and research on attributing negative outcomes to discrimination. As suggested by the ultimate attribution error, the negative outcome (underperformance) of negatively stereotyped outgroup members (Hispanic and Black students) was more likely to be attributed to dispositional and stereotypic causes (Froehlich et al., 2015; Hewstone, 1990; Pettigrew, 1979). And similarly to how Kaiser and Miller (2001) theorized that groups less likely to experience discrimination are more likely to underestimate the occurrence and adverse effects of discrimination, I found that majority students who are presumably less likely to experience stereotype threat were more likely to perceive stereotype threat as a less valid explanation for underperformance. The notion that the perceived validity of an explanation for underperformance is influenced by a person's ability to relate to the explanation and the person providing the explanation is supported by the finding that higher perceived validity was rationalized as more relatable whereas lower perceived validity was rationalized as unlikely to be true.

There are also possible additional reasons for why White nonstereotype-threatened students perceived stereotype threat as a less valid explanation for underperformance compared to test anxiety. One additional reason may be that White students perceived the stereotype threat attribution as less valid because it made issues surrounding race and ethnicity salient in the situation. Research has shown that White Americans generally do not like to discuss issues of race and ethnicity (Bonilla-Silva, 2002). Moreover, attributing underperformance to stereotype threat may be perceived as attributing one's failure to the "potential" prejudice of others. This seems especially true when the source of stereotype threat stems from the concern of being perceived as stereotypic in the eyes of others (Shapiro, 2011). If attributing underperformance to stereotype threat is perceived as blaming others for one's own failures, it may be perceived as breaking the cultural norm of taking responsibility for one's failures (Jellison & Green, 1981). Lastly, any explanation for underperformance provided by a racial/ethnic minority student may be perceived as less valid compared to that of a White student because racial/ethnic minorities are often negatively stereotyped as using their disadvantages (e.g., race/ethnicity) as a means to seek advantage (Wilkins, 2016).

The current research also revealed that the perceived cause for underperformance influenced the feedback that participants provided. Specifically, when no explanation was provided for underperformance (Experiment 1), students whose underperformance was more likely to be attributed to dispositional and stereotypic causes—stereotype-threatened students—were less likely to receive the most helpful form of feedback presumably because such feedback would not change or improve the perceived cause for their underperformance. When an explanation was provided for underperformance (Experiment 2), students were more likely to receive feedback that addressed the provided cause for their underperformance (i.e., anxiety).

These findings suggest that when stereotype-threatened students attribute their underperformance to stereotype threat, nonstereotype-threatened students' feedback is more likely to be influenced by the provided explanation for underperformance rather than their attributions for the stereotype-threatened student's underperformance.

Limitations and Future Directions

One limitation of the current research is that the stereotype threat explanation for the Hispanic and White fellow student was dissimilar. Specifically, fellow students' described as White expressed the concern of being unable to confirm a *positive* stereotype about their racial/ethnic group whereas fellow students' described as Hispanic expressed the concern of confirming a *negative* stereotype about their racial/ethnic group. Thus, it is unclear whether the differences in the perceived validity of providing any explanation for underperformance was influenced more by the race/ethnicity of the fellow student or the differences in the stereotype threat explanation. It is possible that participants perceived explanations provided by Hispanic students as less valid not because of their race/ethnicity but because of the differences in their stereotype threat concerns. However, even if differences found in perceived validity are the result of differences in stereotype threat concerns rather than the race/ethnicity of the fellow student, these differences are simply a reflection of the disparate stereotypes and concerns students may experience as a result of their racial/ethnic identity. Nevertheless, future research should address this issue by examining the perceived validity of attributing underperformance to stereotype threat for different social groups that share an identical threat.

Future research should also examine whether stereotype-threatened students are more likely to detect stereotype threat in others and perceive it as a valid explanation for underperformance. It is possible that stereotype-threatened students are equally unlikely as

nonstereotype-threatened students to perceive stereotype threat as a less valid explanation for underperformance. However, research on collective threat may suggest otherwise (Cohen & Garcia, 2005). Research on collective threat has demonstrated that negatively stereotyped social groups can become concerned that the behavior of an ingroup member may reinforce negative stereotypes about their shared group identity. If students can be consciously aware of their experience of collective threat, they may also be more receptive and understanding to the stereotype threat concerns of a fellow ingroup member.

Final Remarks

Over two decades of research has demonstrated the adverse effects stereotype threat can have on the academic performance of members of stigmatized groups. However, little research has examined how students not experiencing threat perceive students who are, and no research has examined how they perceive stereotype threat as an explanation for underperformance. As more research reveals that stereotype-threatened student can be consciously aware of their experience of stereotype threat, it is becoming increasingly important to examine the possible social consequences and perceived validity of attributing underperformance to stereotype threat. The findings of the current research highlight the need to examine the adverse effects of stereotype threat from the perspective of the unthreatened.

References

- Beilock, S. L., & Willingham, D. T. (2014). Math anxiety: Can teachers help students reduce it? ask the cognitive scientist. *American Educator*, 38(2), 28.
- Bonilla-Silva, E. (2002). The linguistics of color blind racism: How to talk nasty about blacks without sounding “racist”. *Critical Sociology*, 28(1-2), 41-64.
- Boucher, K. L., Rydell, R. J., & Murphy, M. C. (2015). Forecasting the experience of stereotype threat for others. *Journal of Experimental Social Psychology*, 58, 56-62.
- Cheryan, S., & Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance: The psychological hazards of "model minority" status. *Psychological Science*, 11(5), 399-402.
- Cohen, G. L., & Garcia, J. (2005). " I am us": Negative stereotypes as collective threats. *Journal of Personality and Social Psychology*, 89(4), 566.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878.
- Froehlich, L., Martiny, S. E., Deaux, K., & Mok, S. Y. (2016). “It’s their responsibility, not ours” –Stereotypes about competence and causal attributions for immigrants’ academic underperformance. *Social Psychology*, 47,74–86.

- Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of latino women. *Personality and Social Psychology Bulletin*, 28(5), 659-670.
- Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, 29(1), 17-28.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24(6), 645-662.
- Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology*, 20(4), 311-335.
- Jellison, J. M., & Green, J. (1981). A self-presentation approach to the fundamental attribution error: The norm of internality. *Journal of Personality and Social Psychology*, 40(4), 643.
- Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, 16(3), 175-179.
- Kaiser, C. R., & Miller, C. T. (2003). Derogating the victim: The interpersonal consequences of blaming events on discrimination. *Group Processes & Intergroup Relations*, 6(3), 227-237.
- Kaiser, C. R., & Miller, C. T. (2001). Stop complaining! the social costs of making attributions to discrimination. *Personality and Social Psychology Bulletin*, 27(2), 254-263.

- Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology, 44*(4), 645-657.
- Mastro, D. E., & Behm-Morawitz, E. (2005). Latino representation on primetime television. *Journalism & Mass Communication Quarterly, 82*(1), 110-130.
- Niemann, Y. F., Jennings, L., Rozelle, R. M., Baxter, J. C., & Sullivan, E. (1994). Use of free responses and cluster analysis to determine stereotypes of eight groups. *Personality and Social Psychology Bulletin, 20*(4), 379-390.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and social psychology bulletin, 5*(4), 461-476.
- Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current directions in psychological science, 3*(6), 176-179.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt, Rinehart & Winston.
- Shapiro, J. R. (2011). Different groups, different threats: A multi-threat approach to the experience of stereotype threats. *Personality & Social Psychology Bulletin, 37*(4), 464-480.
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology, 67*, 415-437.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology, 69*(5), 797.

Wilkins, C. (2016, January). *Playing the race card: whites believe claiming discrimination is an advantage they don't have*. Talk presented in D. Holoien and C. Wilkins (Chairs), Is that Discrimination? Divergent Perceptions of Discrimination Claims (Symposium) at the 17th annual convention of the Society for Personality and Social Psychology, San Diego, CA

Table 1. Experiment 1: Percentage of overall attributions falling into each category.

Attribution	%
Academic Interests	5.7
Education	20.9
Test Difficulty	18.3
Lack of Effort	21.3
Language	10.9

Note. N =230. Percentages do not reach 100% as some participants' responses were not actual attributions for underperformance and therefore not included in analysis. Moreover, each response could mention more than one attribution category.

Table 2. Experiment 1: Percentage of overall feedback falling into each category.

Feedback	%
Expand Vocabulary	28.0
No Strategy	19.1
Read	23.7
Test-Taking Strategy	37.7
Try Harder	18.6

Note. N =236. Percentages exceed 100% as each response could mention more than one feedback category.

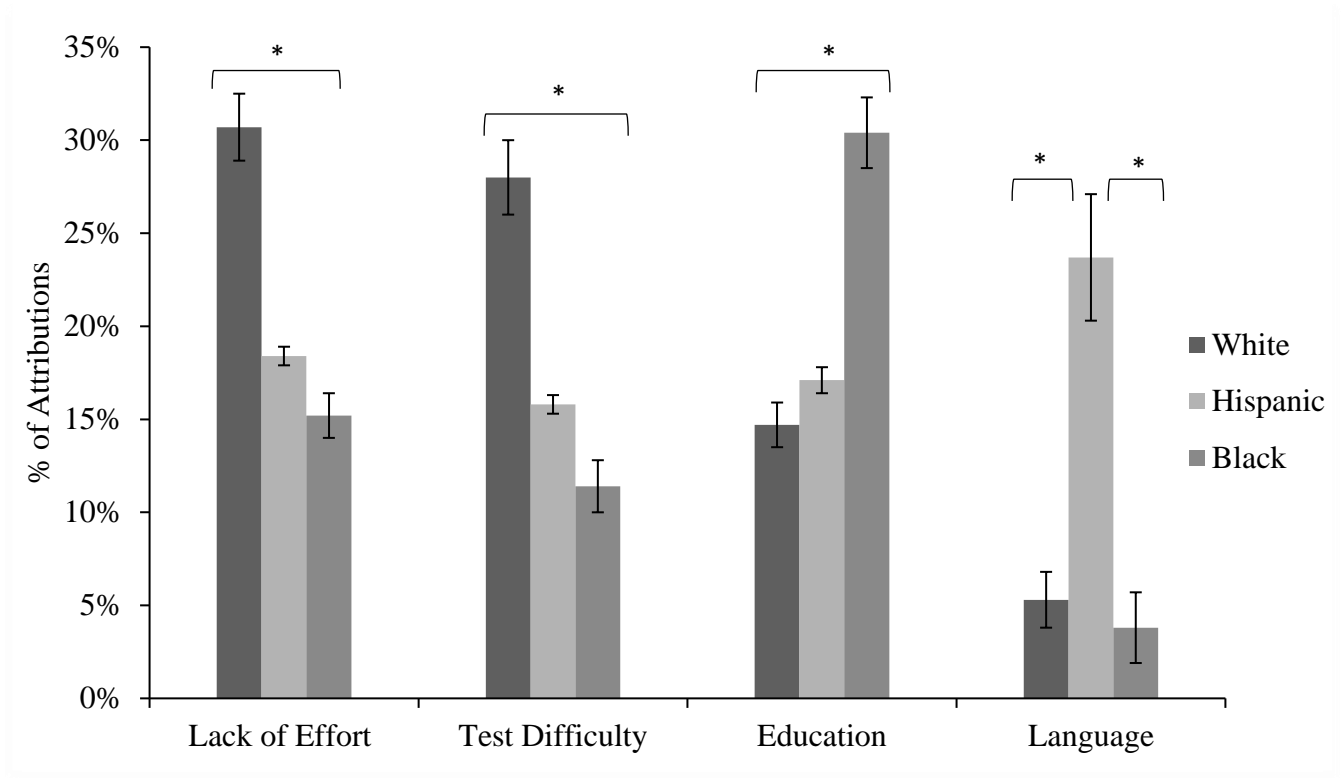


Figure 1. Percentage of attributions provided for underperformance as a function of the fellow student's race/ethnicity in Experiment 1. Error bars represent the standard error. * $p < .05$.

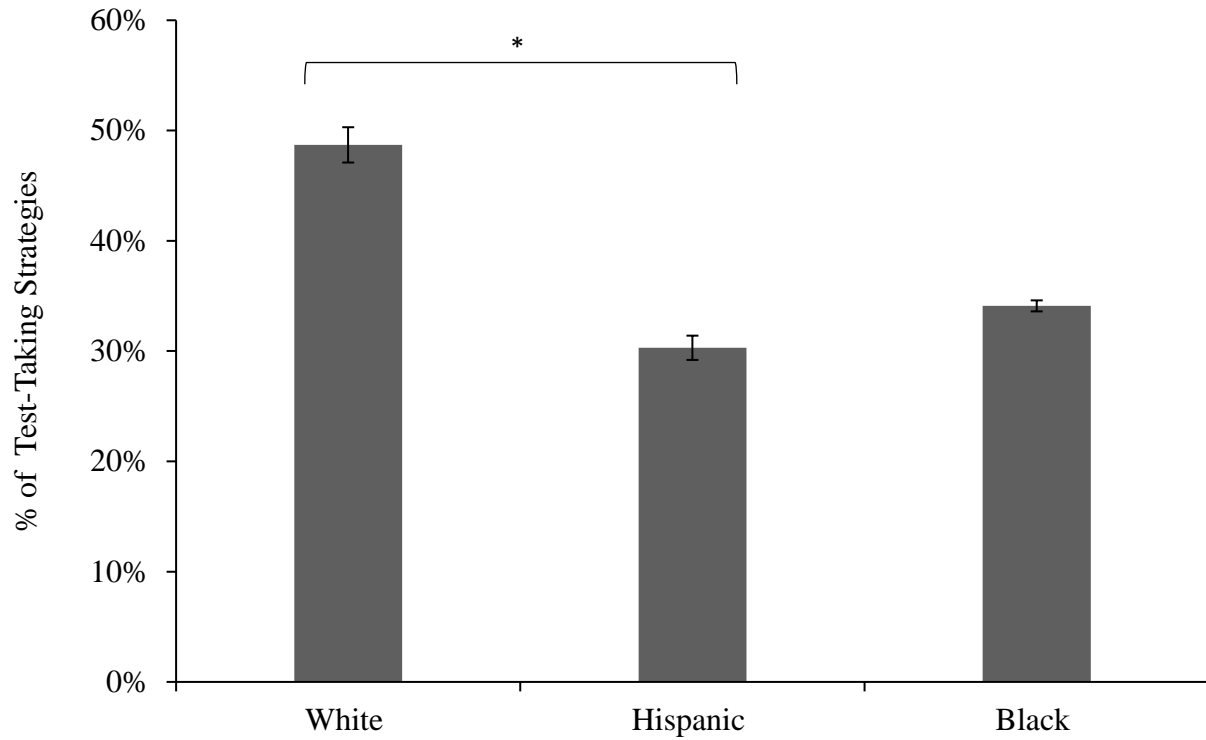


Figure 2. Percentage of feedback provided that was test-taking strategies as a function of the fellow student's race/ethnicity in Experiment 1. Error bars represent standard error. * $p < .05$.

Table 3. Experiment 2: Percentage of overall attributions falling into each category.

Attribution	%
Academic Interests	8.4
Education	20.9
Test Difficulty	13.4
Lack of Effort	19.7
Language	7.1

Note. N =239. Percentages do not reach 100% as some participant's responses were not actual attributions for underperformance and therefore not included in analysis. Moreover, each response could mention more than one attribution category.

Table 4. Experiment 2: Percentage of overall explanations falling into each category.

Feedback	%
Unlikely	37.1
Relatable	22.8

Note. N =239. Percentages do not reach 100% as some participant's responses were not actual explanations and therefore not included in analysis. Moreover, each response could mention more than one explanation.

Table 5. Experiment 2: Percentage of overall feedback falling into each category.

Feedback	%
Anxiety-Reducing Strategies	48.1
Test-Preparation Strategies	38.7

Note. N =243. Percentages do not reach 100% as some feedback provided no strategies and therefore not included in analysis. Moreover, each response could mention more than one feedback category.

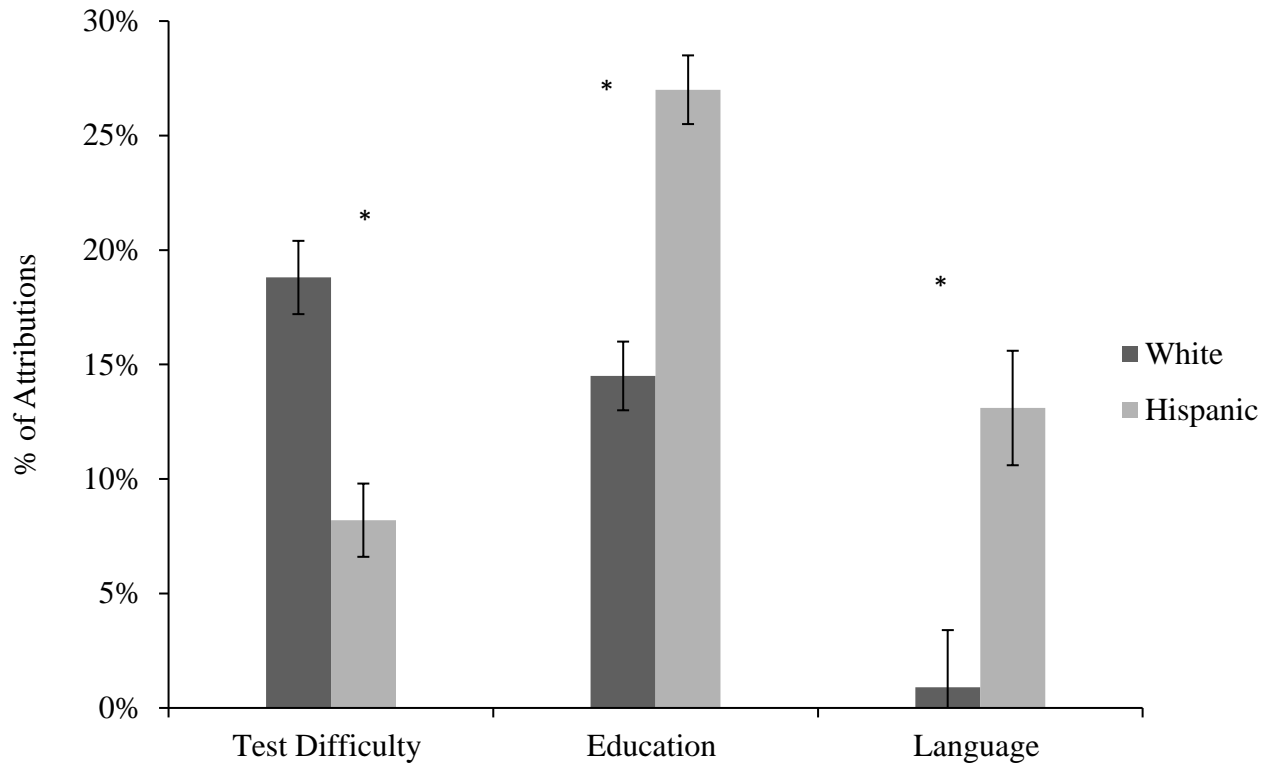


Figure 3. Percentage of attributions provided for underperformance as a function of the fellow student's race/ethnicity in Experiment 2. Error bars represent the standard error. * $p < .05$.

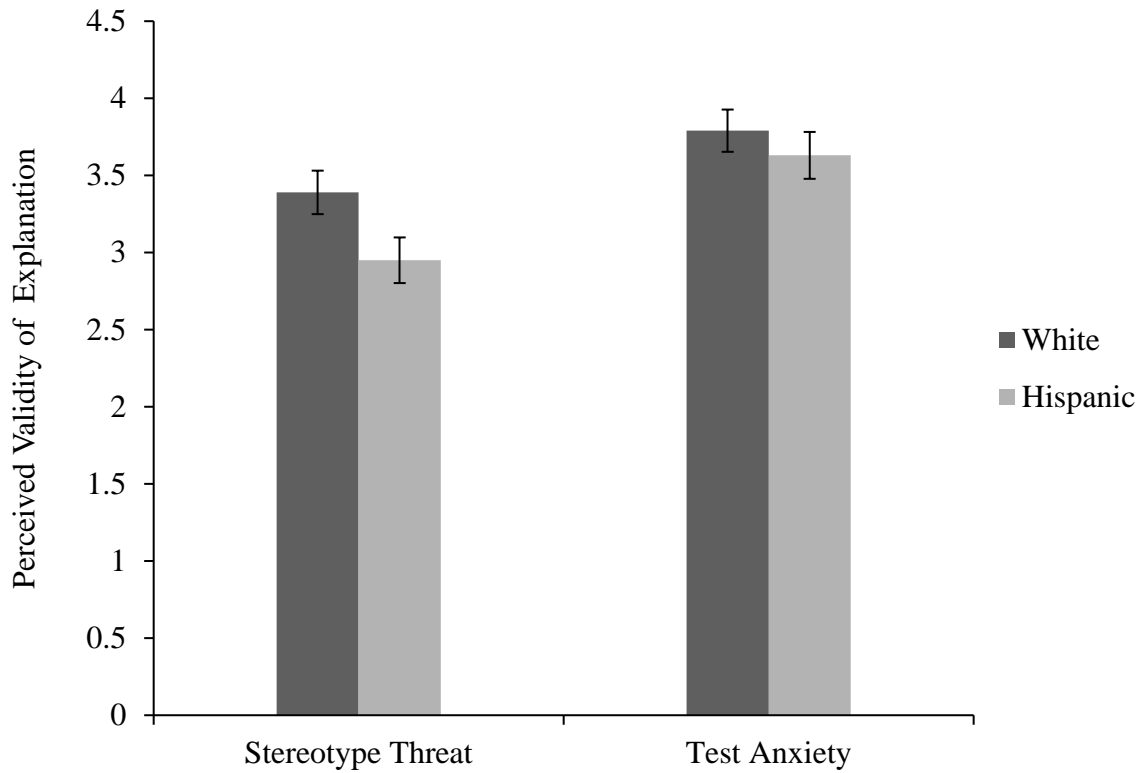


Figure 4. Perceived validity of the provided explanation as a function of the race/ethnicity of the fellow student and the explanation provided for underperformance. Error bars represent standard error.

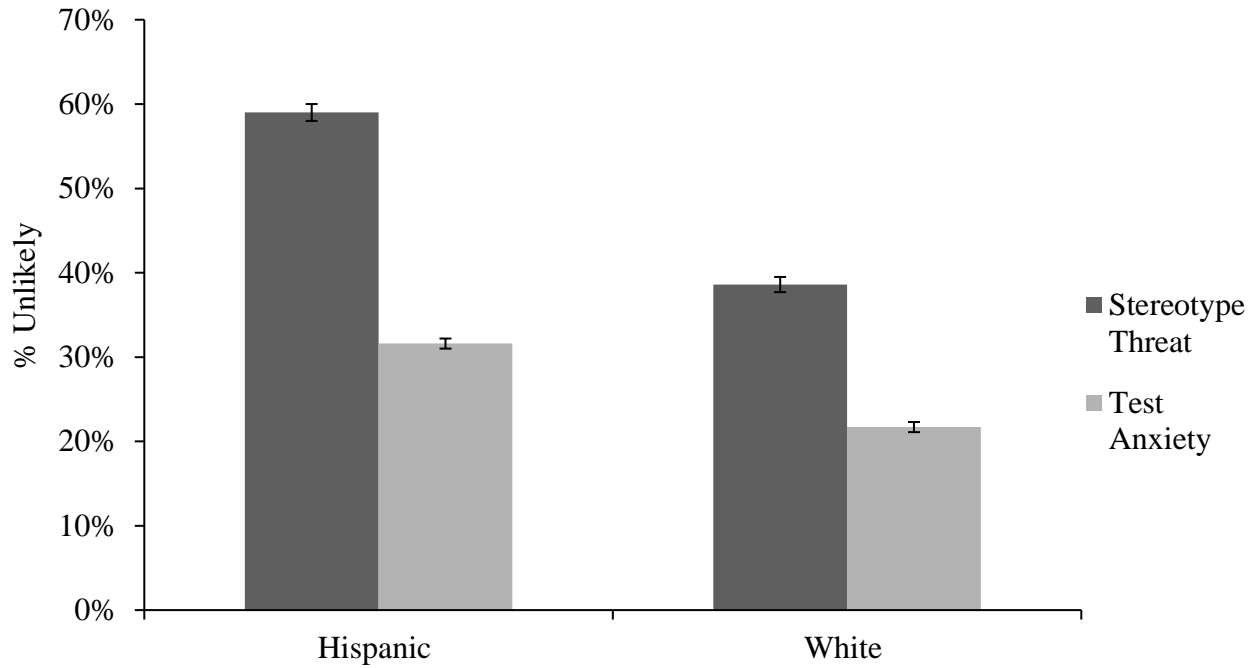


Figure 5. Percentage of responses that stated that the fellow student's explanation for underperformance was unlikely to be true as a function of the fellow student's race/ethnicity and the explanation provided for underperformance in Experiment 2. Error bars represent the standard error.

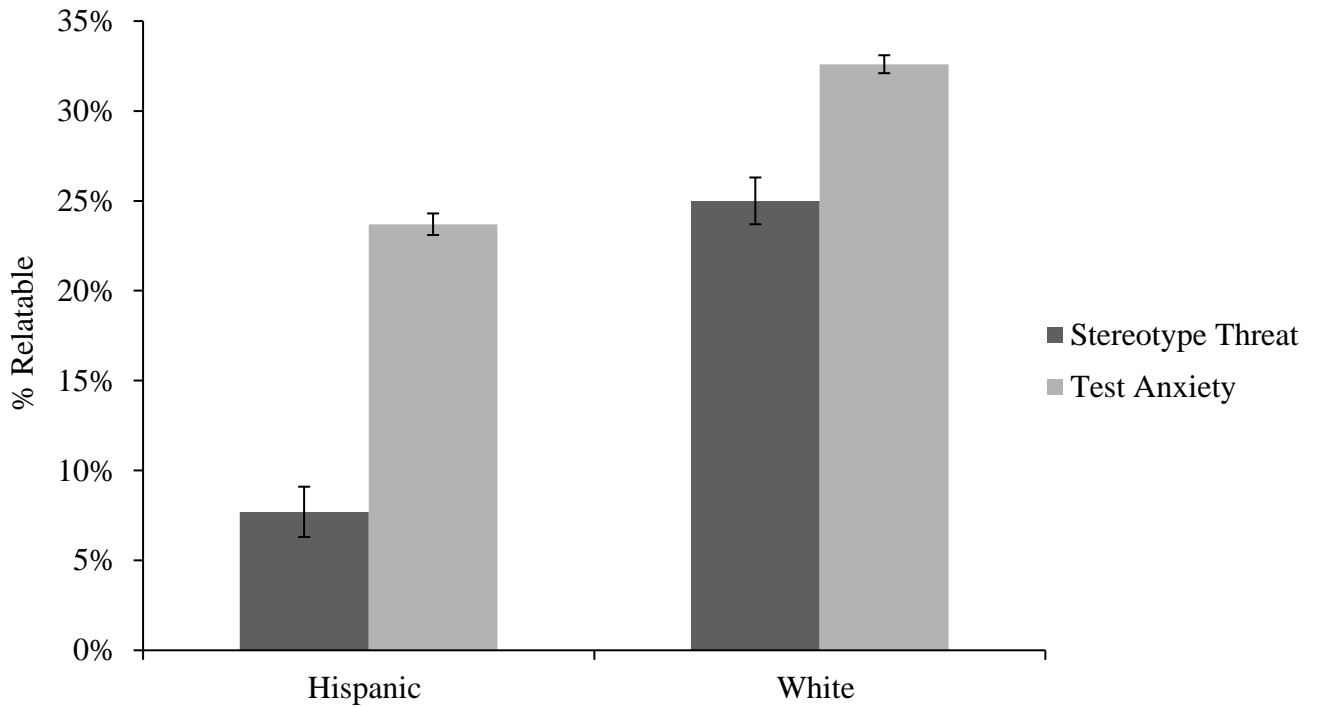


Figure 6. Percentage of responses that stated that the fellow student's explanation for underperformance was relatable as a function of the fellow student's race/ethnicity and the explanation provided for underperformance in Experiment 2. Error bars represent the standard error.

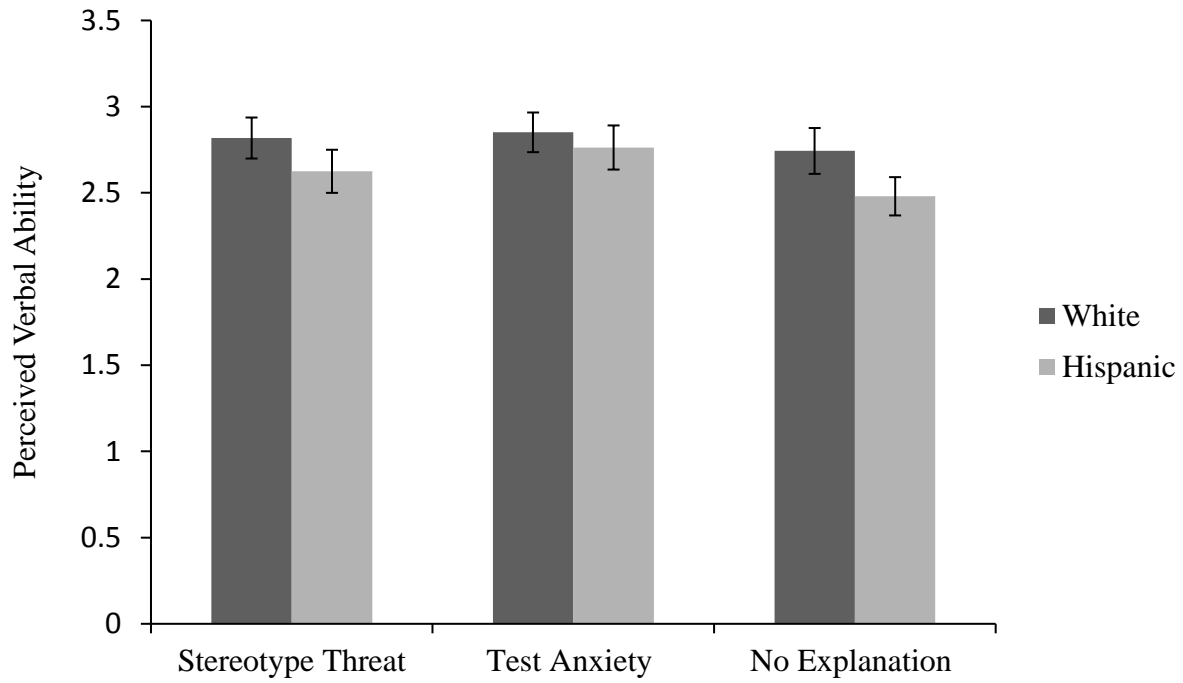


Figure 7. Perceived verbal ability as a function of the race/ethnicity of the fellow student and the explanation provided for underperformance. Error bars represent standard error.

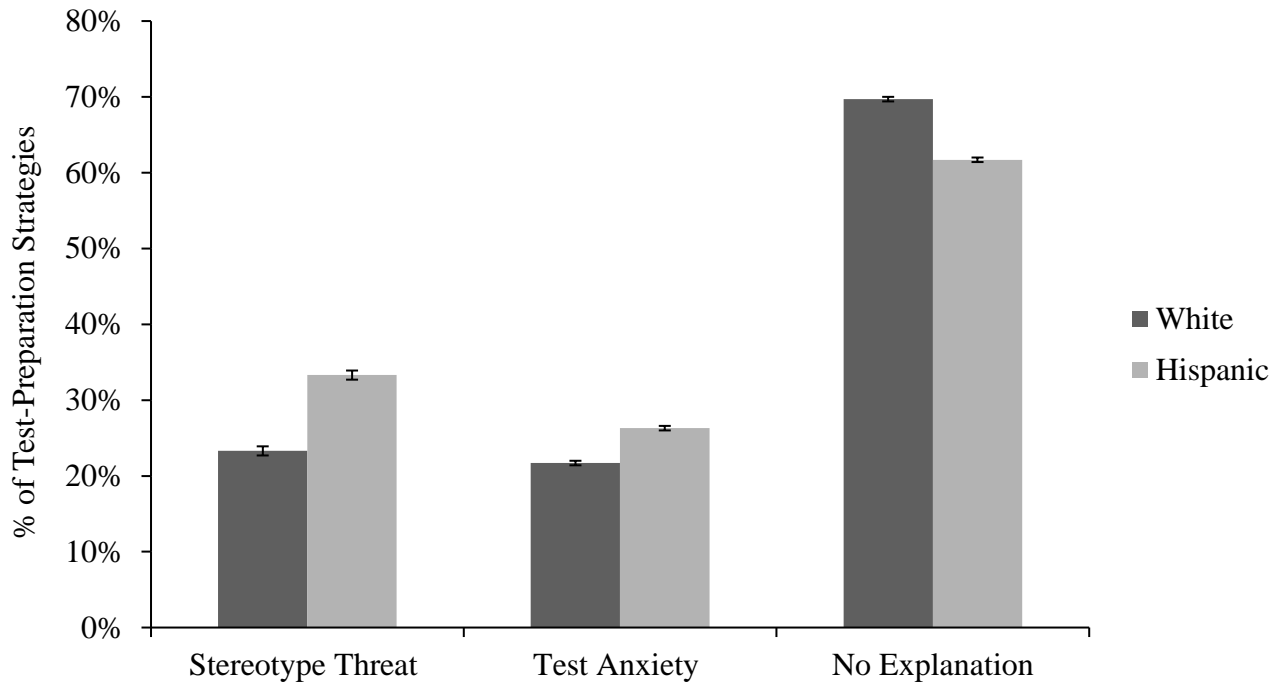


Figure 8. Percentage of feedback provided that was *test-preparation strategies* as a function of the fellow student's race/ethnicity and the explanation provided in Experiment 2. Error bars represent standard error.

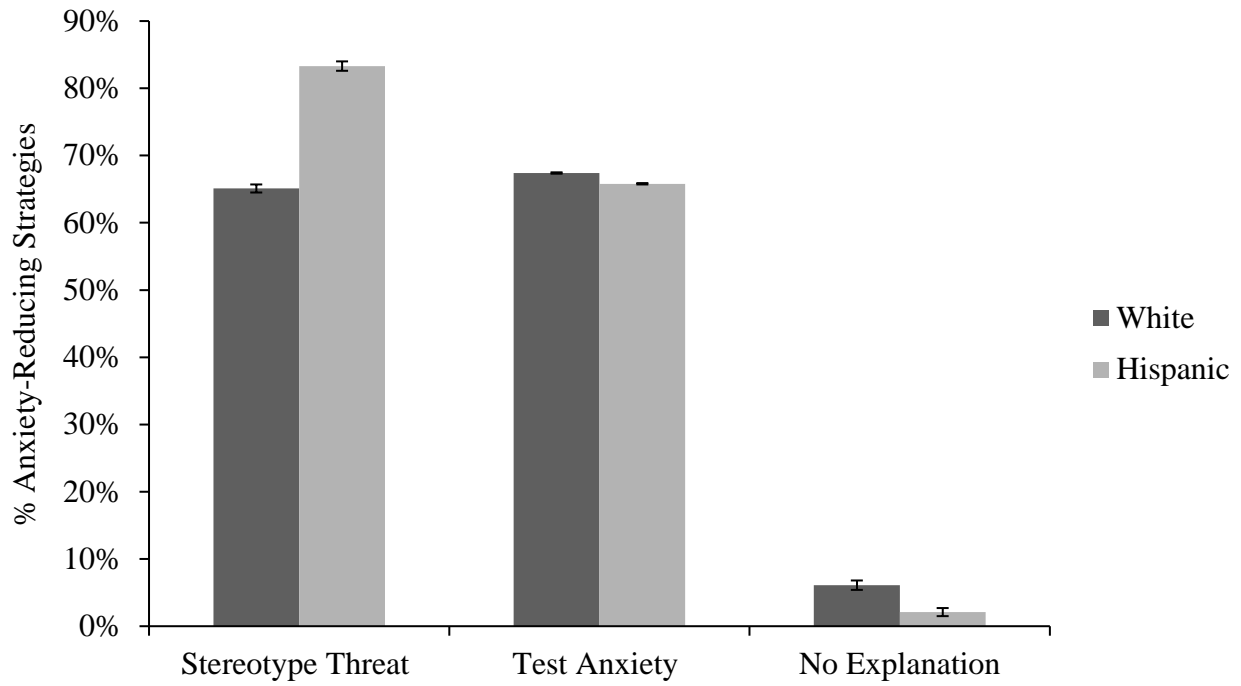


Figure 9. Percentage of feedback provided that was anxiety-reducing strategies as a function of the fellow student's race/ethnicity and the explanation provided in Experiment 2. Error bars represent standard error.