The Report Committee for Leann Vernice Smith
Certifies that this is the approved version of the following report:

Stereotype Threat Vulnerability: A Psychometric Investigation of the Social Identities and Attitudes Scale

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Kevin O. Cokley

Keisha L. Bentley-Edwards

Stereotype Threat Vulnerability: A Psychometric Investigation of the Social Identities and Attitudes Scale

by

Leann Vernice Smith, B.A.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2015

Abstract

Stereotype Threat Vulnerability: A Psychometric Investigation of the Social Identities and Attitudes Scale

Leann Vernice Smith, M.A.

The University of Texas at Austin, 2015

Supervisor: Kevin O. Cokley

Stereotype threat is a widely researched phenomenon within psychology that has been proposed as one explanation for the underperformance of minority groups. Stereotype threat is the experience a person has when negative stereotypes about their social group are highlighted, causing them to underperform on the given task. Picho and Brown (2011) created the Social Identities and Attitudes Scale (SIAS), a scale containing six factors that research has determined to be important moderators of stereotype threat. The current study investigated the psychometric properties of the SIAS. Confirmatory factor analyses and group invariance tests of the SIAS were conducted on a diverse sample of 516 college students participating in a university's subject pool. The results revealed good model fit of the data, with minor exceptions. Additionally, the same factor structure emerged across four different ethnic subgroups: African American, Caucasian American, Hispanic American, and Asian American participants. The SIAS is a reliable and valid measure of six moderators of stereotype threat: ethnic and gender identity, ethnicity and gender stigma consciousness, negative affect, and math identification. Researchers and practitioners can more confidently use the SIAS as a measure of an individual's susceptibility to stereotype threat effects. Future research directions and practical implications are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Educators, policy-makers, and researchers are interested in the academic underperformance of groups of students. In particular, a better understanding of why ethnic minority students are underperforming in academic domains and why women are underperforming in math domains is a frequently discussed topic. As these trends continue, researchers have begun to look for explanations and causes of the disproportionate lack of women and minority students in science, technology, engineering, and math (STEM) domains. Stereotype threat is a common and well-researched source of this underachievement. Therefore, the field is compelled to take a deeper look into the unique experiences of students vulnerable to stereotype threat.

Stereotype threat is a social psychological phenomenon that is characterized by an individual's decreased task performance in the presence of stimuli that increases the individual's awareness of stereotypes about their group's underperformance in the task domain (Steele & Aronson, 1995). Before understanding the scope of the current study, it is important to orient the reader to stereotype threat. To this end, an overview of the stereotype threat's presentation in African American, Latino American, and female populations is provided. A discussion of moderators of stereotype threat follows with the discussion ending on the work of Picho and Brown (2011) and their measure created to operationalize stereotype threat susceptibility.

STEREOTYPE THREAT

Following the work of the authors who discovered stereotype threat, early studies of stereotype threat focused on Black samples. In 1995, Steele and Aronson sought to understand what was causing Black students to underperform when given an intellectual

test to complete. From this curiosity, and multiple experimental studies, emerged the psychological construct of stereotype threat. In their initial publication, which presented the findings of four separate studies, these researchers compared the test performance results of Black and White students. In one of the four studies, students were told that the 30-minute test they were about to take was either a measure of their intellectual ability (stereotype condition) or was just a simple laboratory problem-solving task (control group). Results from their investigation found that in the intellectual task group, Black students performed significantly worse than White students. However, the non-diagnostic group showed that Black and White students performed equally. The researchers speculated that Black students in the diagnostic group believed that the pressure of potentially confirming the negative stereotype of Black people being less intelligent impacted their ability to perform optimally. The additional studies supported this idea by showing that underperformance occurred again when students were presented information that highlighted negative stereotypes about Black performance on intellectual tasks (Steele & Aronson, 1995). This ground-breaking research gave the field a new layer of the achievement gap to explore.

Given that the gap in academic achievement was not solely a Black-White gap, researchers were curious about whether stereotype threat may have an impact on another group experiencing underperformance, namely, Latinos. Research found that Latino students under stereotype threat conditions were also underperforming on tasks. These findings were consistent with those of Black students. Schmader and Johns (2003) designed an experiment to test stereotype threat effects for Latino participants. Latino and Caucasian psychology students, assigned to either a control or stereotype condition group, were told to complete a test that was a reliable measure of working memory capacity. Students in the

stereotype condition group were told that performance on the memory test was highly predictive of intelligence test performance and that results would be used to establish group norms. They were also asked to indicate their ethnicity before beginning the test. Results revealed that Latino students in the condition group recalled fewer words on the memory test than did Latino students in the control group and White students in both groups. Gonzales, Blanton, and Williams (2002) also studied stereotype threat's effect on performance by comparing Latino and White undergraduate students. The researchers tested for the main effect of ethnicity on task performance in both a diagnostic and control groups. Both groups were told that they would complete very difficult problems that would test their math and spatial ability. The diagnostic group was told that the test would be a genuine test of their actual abilities and limitations. Results of this study showed that Latino participants in the diagnostic group scored significantly less than all three of the other participant groups. Additionally, in their meta-analysis comparing stereotype threat effects of Hispanic and African American populations, Nadler and Clark (2011) found that there was no significant difference between the two groups, regardless of study experimental designs and sampling procedures. This wave of research supports extending the stereotype threat construct to all stigmatized ethnic groups, especially Latinos.

Women consistently underperform in comparison to males in math domains after middle school (Huguet & Régner, 2007). Thus, researchers began to investigate the impact of stereotype threat on women and found it to be a possible cause of women's math underperformance trends. To explore whether effects of stereotype threat could be experienced by women, Spencer, Steele, and Quinn (1999) designed a study using men and women with strong math backgrounds and hypothesized that women would score lower than men when the test was mentioned as having gender differences and that they would

score equally otherwise. The sample of undergraduate students was divided into two group that differed only by the order in which gender differences would be highlighted. One group was told that the first section of a computer test had gender differences and the second section did not. The other group was told the reverse. Additionally, all participants were read information about gender differences in math ability. The researchers' hypothesis was confirmed. Women performed lower than equally qualified men when told that the test contained gender differences. Also, in their study on the working memory's role in underperformance under stereotype threat conditions, Schmader and Johns (2003), hypothesized that women would experience lower levels of working memory than men under stereotype threat conditions. Results from their study revealed that women who were told the working memory task was related to math ability showed reduced cognitive capacity, while women in the control condition and men in both the control and condition groups did not. These studies provide evidence that women experience stereotype threat similar to ethnic minorities and that stereotype threat may play a role in the underperformance of women in math.

Studies similar to the ones previously mentioned have been replicated and extended many times, resulting in research studies of stereotype threat in populations not necessarily linked to major educational achievement gaps. Research has found support for stereotype threat's impact on performance outcomes for students in low socioeconomic status, for older adults on working memory tasks, for student athletes on academic achievement, and White males when compared to Asian males in science and math domains (Aronson et al., 1999; Croizet & Claire, 1998; Feltz, Schneider, Hwang, & Skogsberg, 2013; Jameson, Diehl, & Danso, 2007; Mazerolle, Régner, Morisset, Rigalleau, & Huguet, 2012; B. Spencer & Castano, 2007). Although these populations are not being addressed in the

present study, it is important to note the wide range of social domains that can be impacted by stereotype threat.

MODERATORS

To assume that all women and/or members of an ethnic minority group experience stereotype threat would be an over generalization of the experiences of groups of people. Making this assumption would also leave no room for areas of resilience that could allow individuals to overcome the effects of stereotype threat or to not experience the threat at all. Additionally, when conducting research on ethnic minority populations, it is important to move beyond an individual's group status and get to more proximal constructs that involve culture (Betancourt & López, 1993). To this end, researchers set out to get a better understanding of what specific processes were causing the task underperformance of groups experiencing stereotype threat.

According to Baron and Kenny (1986), a moderator is a variable that changes the direction and/or strengthens the relationship between an independent variable and a dependent variable. In this case, researchers began to look deeper into the moderators of stereotype threat's effect on task performance. Given the range of social identities and task domains that have been researched in the stereotype threat literature, it should not be surprising that moderating variables are plentiful. Converging evidence revealed that many processes interact and change the experience of stereotype threat on task performance. The most commonly mentioned moderating variables are domain identification, group identification, stigma consciousness, and locus of control (Brown & Pinel, 2003; Cadinu, Maass, Lombardo, & Frigerio, 2006; Davis, Aronson, & Salinas, 2006; Inzlicht & Kang, 2010; Lawrence, Marks, & Jackson, 2010; Nguyen & Ryan, 2008; Schmader, 2002; Schmader & Johns, 2003; Steele, 1997; Steinberg, Okun, & Aiken, 2012). Although this

list of moderators is not exhaustive, the variables are among the most cited and consistently agreed upon moderators of stereotype threat effects on women and ethnic minorities.

MEASURING STEREOTYPE THREAT VULNERABILITY

To advance the research on stereotype threat, Picho and Brown (2011) created an integrated measure of key moderators of stereotype threat. The authors explain that having a measure of a person's vulnerability to experiencing stereotype threat would make it possible to "establish a baseline for measuring stereotype threat and the impact of interventions attempting to reduce it" (Picho & Brown, 2011, p.377). In their development of the Social Identities and Attitudes Scale (SIAS), Picho and Brown considered the relationships between the constructs that emerged from the stereotype threat literature as being highly influential in a person's experience of the threat. In the scale construction stage of the SIAS development, Picho and Brown decided on seven key constructs and designed the scale based on pre-existing measures of those particular constructs. Given the wealth of research that focused on stereotype threat's impact on math performance, Picho and Brown decided that math should be the domain of interest for the SIAS (however, they assert that the scale can be adapted for other domains). Thus, the seven constructs proposed for the SIAS included math identification, negative affect, math self-concept, gender identification, gender stigma consciousness, ethnic identification, and ethnicity stigma consciousness. Content validity analyses were conducted by ten content validators and results were used to reduce the initially 80-item scale to 43 items receiving an 80% or greater agreement among validators.

The next stage in the development warranted an exploratory factor analysis (EFA) to assess the factor structure of the SIAS. Their EFA of the 43-item, 7-point Likert scale revealed that there were a total of five items that each loaded on more than just it's one,

intended factor. They also found that many items loading on the math self-concept factor were redundant and loaded on other factors as well. Thus, math self-concept and its items were removed from the SIAS along with the other problematic items. A six-factor, 33-item scale was the final result.

To test the psychometric properties of the SIAS, Picho and Brown conducted a confirmatory factor analysis of the scale with a sample of 200 college students. Overall, the results of this analysis upheld the six-factor structure of the scale. To improve fit, the errors of three pairs of items were correlated, three items were deleted, and good reliability estimates and factor correlations were achieved. The final SIAS contains 30 items and has six factors. Factors and sample items can be found below.

The six factors that emerge from the SIAS are listed and defined exactly as Picho and Brown described the factors in their analysis:

Ethnic Identification (EI) is the extent to which an individual forms their identity based on their membership within their ethnic group. The SIAS contains four items that load on the EI factor. One item states, "I am connected with my ethnic heritage."

Ethnicity Stigma Consciousness (ESC) is the extent to which one is chronically self-conscious of stigma attached to one's ethnicity. The SIAS contains five items that load on the ESC factor. One item states, "My ethnicity affects how I interact with people from other ethnicities."

Gender Identification (GI) is the extent to which an individual's gender forms a central part of one's self-concept. The SIAS contains four items that load on the GI factor. A sample item reads, "My gender is central to defining who I am."

Gender Stigma Consciousness (GSC) is the extent to which one is chronically self-conscious of stigma attached to one's gender. The SIAS contains five items that load on the GSC factor. A sample item reads, "My gender affects how people act towards me."

Math Identification (MI) is an identity formed by students who value math, have the skills to succeed in it, and perceive it as being useful to their future career. The SIAS contains six items that load on the MI factor. A sample item reads, "Doing well in math matters to me."

Negative Affect (NA) is associated with negative feelings of dejection experienced during math test taking. The SIAS contains six items that load on the NA factor. All items of this factor begin with the phrase, "When doing difficult problems on a math test I…" A sample item reads, "Feel hopeless".

Strong, positive factor correlations were found between GSC and ESC (See Table 1). This correlation suggests that perhaps there is an overall heightened awareness of stigma that subsumes GSC and ESC. A strong, significant correlation between EI and ESC means that an individual who strongly identifies with their ethnic group is also more conscious of the negative stereotypes that others have about their group. This same positive correlation was found between GI and GSC. The final significant factor correlation was between NA and MI. This moderate, negative correlation suggests that the more one identifies with the math domain, the less that negative feelings associated with math will occur.

Overall, research on stereotype threat has been valuable in helping the field understand some challenges that students may be experiencing which can have a

detrimental impact on their achievement. However, it is important that the knowledge we have on stereotype threat be extended. A psychometrically sound measure of stereotype threat would be helpful in evaluating the effectiveness of interventions by serving as a supplemental assessment of treatment outcomes and as a screener to identify participants that would be more susceptible to stereotype threat. Although the author commends Picho and Brown (2011) for the development of such a comprehensive scale for stereotype threat moderators, the initial factor analyses were limited in their sample diversity. Additionally, no assessment of whether factor structures differed across different groups was conducted. It is important and recommended by Okazaki and Sue (1995) that all measures that will be used on ethnic minority groups be tested for equivalence. Thus, the purpose of this study is to psychometrically support the development of a comprehensive measure of an individual's susceptibility to stereotype threat as well as offer suggestions for measure improvement based on the findings of a confirmatory factor analysis and group invariance test.

THE PRESENT STUDY

Both researchers and practitioners depend on good measures to collect data and better understand clients. Given the importance of measures, it is fitting that they go through psychometric evaluations with different populations to support their reliability and validity. Although the developers of the SIAS did a great job in measure development and assessing the reliability and validity, it is important that additional psychometric research confirm that the model fit the data for more heterogeneous ethnic groups since these groups would be expected to be more susceptible to stereotype threat. It is anticipated that a measure of an individual's stereotype threat vulnerability would contribute to both practice and research. Practitioners rely on researchers to develop comprehensive, yet concise,

measures that can be used to aide in diagnosis and treatment planning. Because the SIAS provides brief measures of six constructs that all have their own, separate instruments and are all key to stereotype threat, clinicians can use the SIAS to quickly get a measure of a client's stereotype threat vulnerability. This could ultimately impact the clinician's treatment plan or be paired with their subjective evaluation of the client's presenting concern. In addition to academic underperformance, stereotype threat may help explain anxiety, reduced cognitive capacity, academic and domain disidentification, distancing oneself from the stereotyped group, and altered professional identities (Marx & Staple, 2006; Schmader & Johns, 2003; Steele & Aronson, 1995; Steele, James, & Barnett, 2002; Woodcock, Hernandez, Estrada, & Schultz, 2012). Research on stereotype threat has shifted gears and now seeks to develop interventions and ways to reduce stereotype threat's impact on performance. The SIAS can help advance that research by being included in experiments as screeners that may help place students in treatment and control groups, as well as a measurement of intervention effectiveness. However, if this measure is to be used in future research and practice, reliability and validity must be further investigated.

The current study seeks to aid in the continuous validation process by evaluating the SIAS on a more diverse sample. The sample used in this study is more heterogeneous than the sample used in the initial validation of the SIAS. This added diversity could lead to support for, or evidence against, generalizing the SIAS to other populations. Table 2 shows the participant characteristics of Picho and Brown's analysis and of the sample used in the current study.

Overall, the main purpose of this study is to investigate the psychometric properties of Picho and Brown's measure of key moderators of stereotype threat, the SIAS. In particular, the goals of the current study are to (1) measure the internal consistency of the

SIAS, (2) test if the responses of the diverse sample used in the current study generate the same six-factor model as previously found, (3) ensure that the same factor structure emerges across four different ethnic groups, (4) and determine if there are any indicators of measure improvement. Based on the findings of over 300 studies on stereotype threat, most of which included ethnic minorities and women, it is not anticipated that the factor structure of the SIAS would change with a more diverse sample. Additionally, there is no rationale for the model fit and factor structure to differ for different ethnic groups.

Chapter 2: Method

PARTICIPANTS

The participants for this study were 516 college students enrolled in an undergraduate psychology course at a large southern university. The sample included 153 males and 363 female students. Participants self-identified their race/ethnicity. The sample is 21.5% African American, 24.2% European American, 20% Asian American, 22.5% Hispanic, 7.6% Biracial and 4.2% "others". The classification of students was 9.7% freshman, 18.4% sophomore, 24.2% junior, and 46.7% senior.

INSTRUMENT

The Social Identities and Attitudes Scale (Picho & Brown, 2011) is a 30-item measure of an individual's stereotype threat susceptibility based on six constructs that previous research has found to be key moderators of stereotype threat's impact on performance. Previous investigations found that six factors emerge from the scale: Ethnic Identity (EI; 4 items), Gender Identity (GI; 4 items), Ethnic Stigma Consciousness (ESC; 5 items), Gender Stigma Consciousness (GSC; 5 items), Negative Affect (NA; 6 items), and Math Identification (MID; 6 items). All items used a 7-point Likert scale (1= *strongly disagree* to 7= *strongly agree*). An individual's response to the SIAS is interpreted as six different factor scores rather than as an overall score. In Picho and Brown's (2011) development of the SIAS, the researchers found support for discriminant and convergent validity of the measure. Their reliability analysis revealed that the SIAS is a reliable

measure with factors' alpha values ranging from .81 to .95. Specific alpha values for each factor can be found in Table 1.

PROCEDURE

The primary investigator gained access to the university subject pool through the Department of Educational Psychology (EDP). The subject pool is comprised of undergraduate students that are enrolled in an EDP class that offers course credit for research participation. The subject pool is monitored by the university's Institutional Review Board (IRB) and researchers interested in utilizing the subject pool must have their study reviewed by the IRB. The primary investigator (PI) received IRB approval for the current study and was granted access to the subject pool. Once the PI gained access to participant contact information, an electronic link was emailed that allowed the students to participate in the study.

Students enrolled in undergraduate educational psychology courses are given the option to participate in research through the university subject pool. A portion of the student's grade is contingent on their involvement with research. Students are given the option to either participate in the university subject pool as a research study participant or to write a research paper. Once students decide to enter the subject pool, the subject pool research assistant assigns participants to specific research studies based on the student's responses to a few demographic questions and the study's targeted participant population. Students assigned to participate in the current study were sent questionnaire packets electronically that contained the SIAS as well as a demographic information form.

Students were informed that they were involved in a study that seeks to understand social identities and academic outcomes. Students were informed that completion of the electronic survey would be equivalent to them agreeing to participate in the study.

Students completed the SIAS via Qualtrics, a survey-building website. Because the data collected were used to give student's completion grades in their class, some students took the survey more than once. After multiple responses were accounted for and before analyses were conducted, student identifications that were used to assign course credit were removed by the primary investigator to maintain the anonymity of respondents. The primary investigator simply removed that column from the data set after assigning course credit to participants. Students were informed that their student identification numbers would be used for course requirements and were assured that they would not be linked to specific information provided for the study.

STATISTICAL ANALYSES

A series of statistical analyses were conducted to achieve the four goals of this study. First, Cronbach's alpha values were calculated to assess the reliability of the SIAS. Next, a confirmatory factor analysis was conducted on the total sample of 516 college students to assess the factor structure of the SIAS. Following the CFA, change in Chi-Squared tests were conducted to test a series of nested models for model modifications as well as part of the tests of group invariance.

To answer the first research question of the current study, internal consistency was assessed by conducting inter-item correlations among items of each factor. This

analysis gives support of the reliability among the six groupings of items based on the six different factors. More specifically, items that are correlated and load on the same factor show that the items are measuring a common construct. Following the reliability analysis, Confirmatory Factor Analyses (CFA) and tests of group invariance of the SIAS were conducted using Analysis of Moment Structures (AMOS) software. AMOS is a structural equation modeling and path analysis add-on module for SPSS. The initial CFA was conducted on the total sample of 516 college students to assess the factor structure of the SIAS. Following the CFA, tests of change in Chi-Squared values were conducted to test competing, nested models. These tests help to identify potential areas of model or measure improvement. Once a suitable baseline model was decided on, group invariance tests were used to either confirm or reject the hypothesis that the same factors emerge for different ethnic groups. More details on each stage of the analyses follow.

Cronbach's alpha values indicate the internal consistency of an instrument. Alpha values range from 0 to 1. Larger values indicate that the items correlate in a way that will collectively measure a given construct. Cronbach's alpha is a function of inter-item correlations and the number of items measuring the given factor. According to Peterson's (1994) meta-analysis on Cronbach's alpha values, an alpha of at least 0.7 is evidence of internal reliability of the items on a particular scale. A Cronbach's alpha value cutoff of 0.8 was determined to be consistent with Picho and Brown's chosen cutoff value while also satisfying Peterson's recommendation by exceeding his minimum recommended value. Alphas below 0.8 were considered to either have an insufficient number of items

or poor inter-item correlations amongst the items. Given that Cronbach's alpha tests the internal consistency of an instrument, alpha values obtained from different samples should not differ significantly to maintain that the items are stable measures of the underlying construct. A comparison of alpha values from the previous factor analysis of the SIAS and values obtained from this study was made. As part of the reliability analysis output, SPSS reports how deletion of a particular item would affect the factor's overall alpha value. Considering that measure improvement is one of the goals of this study, alpha values if item were deleted were considered.

Multiple model fit indices were used to assess overall model fit to the data during the confirmatory factor analysis. According to Hu and Bentler (1998, 1999), different indices are sensitive to sample size and different aspects of fit. It is recommended to use various indices of comparison to ensure the model fits. For this study, the Root Mean Square Error of Approximation (RMSEA) was used because it is regarded as a very informative fit statistic and allows for the use of confidence intervals (MacCallum, Browne, & Sugawara, 1996). Smaller RMSEA values are desired and a cutoff of 0.06 was used to determine adequate model fit (Hu & Bentler, 1999).

A test of Chi-Square is traditionally used to evaluate model fit. Although it is recommended to report in all factor analyses, Chi-Squared values are very sensitive to large sample sizes. In particular, it almost always reports a significant statistic when the sample size is large suggesting that the model has poor fit to the data. Some researchers have recommended using a normed Chi-Square instead of the regular Chi-Square statistic

when using large sample sizes (Wheaton, Muthén, Alwin, & Summers, 1977). This normed Chi-Square value is calculated by dividing the Chi-Square by the degrees of freedom. Both the regular Chi-square and normed chi-square statistic were used in this study. The normed Chi-Square cutoff ratio used was a value of 3:1 with lower ratios indicating good model fit (Kline, 2010).

The final absolute fit index used was the Standardized Root Mean Square Residual (SRMR) in determining the overall model fit. SRMR values of 0.08 (Hu & Bentler, 1999) have been said to indicate adequate model fit while values less than 0.05 are most desirable (Byrne, 2013). In addition to the absolute fit indices, an incremental fit index was used as is recommended by Kline (2010). The Comparative Fit Index (CFI) was used because it takes sample size into account and is one of the most used fit indices. CFI values range from 0 to 1.0 with larger values indicating better fit. According to Hu & Bentler (1999), values greater than 0.95 indicate good model fit to the data. In addition to the CFI, the Tucker Lewis Index was also reported when comparing models and values close to .90 indicate adequate model fit.

A test of the change in Chi-Square ($\Delta\chi^2$) was used to compare competing, nested models as well as testing for group invariance. A significant result from the $\Delta\chi^2$ test suggests that the models fit the data differently and the model with the most freely estimated parameters should be chosen as the best fitting model. Similarly, if the $\Delta\chi^2$ is not significant the two models do not differ and the model with the least amount of parameters freely estimated should be chosen.

Chapter 3: Results

The goals of this study were to measure the internal consistency of the SIAS, test if the previously found six-factor model was maintained with this diverse sample, test if the factor structure was consistent across ethnic groups, and determine if there were any areas of measure improvement. All six factors of the SIAS yielded sufficient values of Cronbach's alpha values ranging from 0.81 to 0.94 (See Table 1). The alpha values between the two studies were consistent, thus indicating strong support for the reliability of the scale. As mentioned, alpha values of each individual item if deleted were monitored. Of the 30 items, only one item revealed that deletion would not negatively impact the overall alpha value. In particular, deletion of item 25, which loaded on the NA factor, would result in an overall alpha value of 0.93 which is the same alpha value if the item is kept. Given that the overall alpha values did not differ from those found in Picho and Brown's (2011) analyses, it was determined that the SIAS items were reliable; thus further analyses were performed.

Factor correlations were observed to ensure that correlations were consistent with previous findings (See Table 1). Picho and Brown (2011) found significant correlations between each identity factor and their respective stigma consciousness. They also found a positive correlation between both sigma consciousness factors and a negative correlation between negative affect and math identification. The same factor correlations were replicated in this study's results. In addition to the previously found correlations, social

identity factors were more correlated in this sample. Specifically, a positive correlation was found between EI and GI, EI and GSC, and GI and GSC.

To answer the primary research question of whether the same factor structure used in the original validation of the SIAS measure fits the data, a confirmatory factor analysis on the entire sample of 516 participants was conducted. The results of the first analysis are shown in Figure 1. An evaluation of the model fit indices revealed moderate model fit of the data. The analysis generated a χ^2 of 1346.188 (390), p<.001, CFI=.914, RMSEA=.069, SRMR= .046. Although the SRMR and RMSEA values suggest adequate fit based on decided cutoff values, further investigation of whether model fit could be improved revealed that correlating errors may enhance overall model fit to the data.

MODEL MODIFICATIONS

Although modifying a nested model can reveal that additional factors that have not been accounted for are emerging from the data, correlating errors that load on the same factor is a safe way to ensure that this is not happening in the data. A total of six parameters were freely estimated in this model. This may seem excessive; however, Byrne (2013) says that as long as there is a strong reason for errors to be correlated, correlating items loaded on the same factor poses no threat to fit. The tests of change in chi-square values for each freed parameter can be found in Table 3 along with the fit statistics for the initial model.

Item 25 and Item 27. Both items 25 and 27 load on the Negative Affect factor. A look at the modification indices revealed that the largest index value was 106.68 and would result in a parameter change of 0.583. A closer look at the items revealed that both items dealt with doubting one's math abilities. Perhaps an overall lack of confidence may explain the variance between these two items. The researcher decided that based on the modification indices and relationship between the items, the correlation between the two of them should be freely estimated. The overall $\Delta\chi^2$ was significant, $p<.001$, and indicated that freely estimating the correlation between the errors resulted in a better fitting model.

Item 26 and Item 27. Observing the overall model fit indices after correlating the errors of items 25 and 27 indicated the model still did not fit optimally. Thus, further investigation about ways to improve fit reveal another large modification index. The index between item 26 and 27 was 58.9 and correlating their errors could result in a parameter change of 0.358. Both items load on the Negative Affect factor. Item 26 describes feeling like one has let themselves down while completing math problems. The other item, as mentioned previously, deals with a loss of confidence while completing math problems. Perhaps self-efficacy could explain some of the variance in these items. Self-efficacy is thought to include self-esteem and locus of control which appear to both be at play in these items (Judge, Erez, Bono, & Thoresen, 2002). The researcher decided that based on the modification indices and relationship between the items, the correlation between the two of them should be freely estimated. The overall $\Delta\chi^2$ was significant, $p<.001$, and indicated that freely estimating the correlation between the errors resulted in

a better fitting model. Additionally, these are two item errors that Picho and Brown allowed to be freely estimated; thus, correlating them is consistent with previous research.

Item 7 and Item 10. Another look at modification indices revealed that correlating the errors of two items that load on the Gender Identification factor would result in a parameter change of 0.431. Items 7 and 10 actually seem to be asking the same question. One states "my gender is central in defining who I am" and the other says "my identity is strongly tied to my gender". Considering how similar these two items are to one another it makes sense to allow a correlation to be freely estimated. This added correlation resulted in a significant change in model fit, $p<.001$; thus, the correlated model was maintained.

Item 2 and Item 21. Both items 2 and 21 load on the Math Identification factor. Similar to the previous two items that were allowed to have correlated errors, these two items seem to be making the same statement. Item 2 says that "math is important to me" while item 21 says "I value math". To value something means that it is important to you; thus, it makes sense for these two items to be related beyond what would be expected for a person who identifies with the math domain. A significant change in model fit, $p<.001$, supported the correlation of these two items' errors. Additionally, these are two item errors that Picho and Brown allowed to be freely estimated; thus, correlating them is consistent with previous research.

Items 12 and 15; 25 and 26. The last two correlated error terms were made consecutively after item observations revealed that a similar trend of the items asking the same question was occurring. Both correlations resulted in a significant change in model fit. Additionally, item errors for number 12 and 15 are consistent with errors freely estimated in Picho and Brown's factor analysis. Figure 2 represents the final model used and includes a total of six added correlation of error terms. The final model achieved a χ^2 of 927.00(384), $p<.01$, CFI=.95, TLI=.945, RMSEA=.052 which indicates good model fit has been established.

GROUP INVARIANCE TEST

The group invariance test seeks to confirm whether construct validity is maintained. The sample of 516 students was split to represent four ethnic groups: European American, African American, Asian American, and Hispanic American. Students that self-identified ethnically as 'biracial' or 'other' were excluded from this test of group invariance to maintain the integrity of their identity and also to prevent an over generalization of the experience of biracial and international students. This resulted in group sizes of 111 African Americans, 125 European Americans, 103 Asian Americans, and 116 Hispanic Americans.

A test of group invariance requires a hierarchy of steps in which parameters are progressively constrained to be equal across groups. The first step is to test the configural model. This model serves as a baseline by which future models' fit statistics will be compared. The fit of the configural model is tested to ensure that the number of factors

and the pattern of the structure were similar across all groups. The model used was the final model which included the six freely estimated error covariances. The model generated a χ^2 of 2642.8(1544), p<.001, CFI=.892, TLI=.878, RMSEA=.040. These statistics reveal adequate fit; thus, factor structure is similar across groups. The next step is to ensure that the measurement model is the same across groups. This tests whether the factor loadings are equal across all groups. The measurement model with factor loadings constrained to be equal across groups generated a χ^2 of 2747.2(1628), p<.001, CFI=.890, TLI=.882, RMSEA=.039. A significant change in chi square will tell us that the less constrained model (in this case, the configural model) is retained. This would mean that factor loadings do differ across groups and further investigations would be required. However, the test of change in chi-squared values was non-significant, $\Delta\chi^2$ of 104.4(84), p=.07; thus, the measurement model is retained and factor structure and loadings are invariant across the four ethnic groups.

Chapter 4: Discussion

An important goal of this study was to conduct a confirmatory factor analysis of the SIAS to ensure that the six-factor structure of the measure was supported. Based on the model fit statistics of the initial model with uncorrelated error terms, the six-factor structure of the SIAS was adequately supported. A series of correlated error terms revealed that model fit could be improved significantly, however, providing stronger support of the measure's model. These findings indicate that the SIAS is a reliable and valid measurement of an individual's level of the six key constructs that research has found to moderate the impact of stereotype threat on performance: ethnic and gender identity, ethnic and gender stigma consciousness, negative affect, and math identification.

The primary goal and most notable contribution of the current study was to test if the six factor structure that previously emerged in the initial development of the SIAS was consistent across different ethnic groups. Although this step is often left out in measure development and validation, measurement invariance must be maintained if we anticipate that the measure will be used for different racial and gender groups (Horn & Mcardle, 1992). The test of invariance in this study revealed that the same factors emerge across samples of Black, White, Asian, and Hispanic college students. This finding further supports the use of the SIAS in potentially identifying stereotype threat susceptible individuals, regardless of their racial/ethnic status. Given that stereotype threat is a phenomenon that primarily affects individuals belonging to stigmatized groups, support for group invariance was a necessary step in the scale development process.

Because the measure is invariant, researchers and practitioners can give this measure to individuals belonging to any ethnic group and can be confident that it is still measuring the six factors that the SIAS claims to measure.

The correlation of error terms leads to a few suggestions for potential measure improvement. Three of the items that load on the Negative Affect factor, namely 25, 26, and 27, all ended up having correlated error terms. Given that these three items all had the same questions stem; specifically, "When doing difficult math problems on a test I…," it is anticipated that there will be a high correlation. However, correlated errors of the other three items that load on this factor and have the same stem did not emerge as a suggested modification suggesting that there is an underlying similarity between those three items over and beyond that of the question stem and the factor. Also, recall that deletion of item 25 will not affect the Cronbach's alpha according to the reliability analysis. Perhaps these three items are measuring a construct of math self-efficacy or self-esteem. Condensing these three items into a single item may shorten the length of the measure and prevent the addition of three correlated item errors. This gives additional support for removing the item. Similarly, the three other correlated errors in the final model may suggest redundancy of items. Perhaps combining these correlated items would also reduce measure length without affecting the psychometric properties of the SIAS.

LIMITATIONS

Two notable limitations of this study should be discussed. The first short-coming is that the test of group invariance omitted respondents that endorsed their ethnic status as other and biracial. The number of students in these categories was not large enough to maintain enough power that would lead to interpretable results. Also, the researcher found it unethical to group all of these students in the same group because of the anticipated diversity and uniqueness of each of their experiences as a member of their particular racial group(s). Evidence of how well this measure works for multi-ethnic students and international students would add a level of validity that is necessary in the measure development process. Future research should explore this measure with those populations.

The other limitation of this study was that the sample only included college students at a predominantly White institution, similar to the initial validation of the SIAS. This limits the researcher's ability to generalize the measure's utility for non-college samples. Additionally, it could be argued that students who are in college may represent a group of individuals that are resilient when faced with stereotype threat. Since one of the primary uses of this measure is to identify students who may be vulnerable to stereotype threat to create interventions to promote academic achievement, future research should seek to validate this measure on middle or high school samples.

IMPLICATIONS

The Social Identities and Attitudes Scale is unique in that it is the first scale designed that includes six key, research-supported moderators of stereotype threat. This

scale is an important addition to research as our curiosity about stereotype threat shifts to practical implications and intervention development. Many of the interventions being created to buffer the deleterious effects of stereotype threat on task performance have focused solely on status endorsement rather than status identity. Because this measure includes factors that are considered more proximal to the mechanism by which stereotype threat causes underperformance (e.g., anxiety and reduced working memory), it can be viewed as an efficient and more culturally appropriate way to capture an individual's susceptibility to stereotype threat effect.

Despite study limitations, this study adequately investigated the psychometric properties of the Social Identities and Attitudes Scale, which is necessary for the ongoing process of instrument development. This study confirmed that this measure is both reliable and valid and that the underlying factors are the same for various ethnic groups. This finding should be viewed as an important addition to the stereotype threat literature and has the potential to help the field to further research efforts aimed at promoting women and ethnic minority's success in STEM domains. This measure can be used by both practitioners and researchers. It has been recommended that school psychologists consider stereotype threat when performing psycho-educational testing and also when providing treatment for academic-related anxiety (Jordan & Lovett, 2007). Individuals who have begun to design interventions for low-achieving students can use this measure as a way to track student progress. Additionally, researchers can create more experiments

by including the SIAS as a screener for treatment group assignment and also as an outcome measure.

Researchers have already begun to investigate the ways that membership in more than one stereotyped group may change the degree of stereotype threat that an individual experiences (Gonzales et al., 2002; Shapiro, 2011). For example, given what we know about stereotype threat, ethnic minority women could potentially be at a higher risk of stereotype threat than an ethnic minority male in math domains. The SIAS can be used to categorize individuals as having high, low, or moderate levels of stereotype threat vulnerability as well as identifying which particular factors of stereotype threat vulnerability are most salient. This can be beneficial in selecting the way to intervene, as explained in Shapiro et al.'s (2013) development of the multi-threat approach to stereotype threat intervention. The researchers purport that it is most effective to design an intervention based on the range of possible stereotypes that could be experienced by each individual.

It is important to note that, as with all measurement instruments, the SIAS requires further psychometric investigations. Future investigations should focus on assessing the psychometric properties of the SIAS on non-college samples. If this measure is to be used to identify individuals that are more susceptible to stereotype threat, it will be important to ensure it can be utilized on adolescent samples of middle and high school students. This will become increasingly important as the racial and

gender gap of STEM achievement continues to affect younger cohorts of middle school students.

The creators of the SIAS decided to construct the scale using math as the domain of interest, limiting the utilization of the scale to persons interested in identifying individuals susceptible to math stereotype threat. The researchers note that the scale can be modified to include domains other than math. Given that stereotype threat literature is not limited to math domains, being able to use this scale for other areas of performance is important. Future scale investigations should explore the psychometric properties of the scale in domains other than math such as reading or overall academic identification.

While substantial gains have been made in the scores of women and minorities in some areas of STEM education, there is still a notable gap in the retention of women and minorities in STEM professions. Since research has shown that stereotype threat may cause individuals effected to de-identify with the particular domain, it can be assumed that stereotype threat may be one of the reasons of this exit from stem domains (Woodcock, Hernandez, Estrada, & Schultz, 2012). If the field hopes to see that these gaps get closed, more rigorous research on overcoming stereotype threat becomes imperative. The Social Identities and Attitudes Scale, along with experiments and intervention research on larger, diverse samples, can serve as a step in the problem-solving direction.

Appendix A

Table 1

Previous and Current Correlations of the Six Factors of the Social Identities and Attitudes Scale (SIAS), Alpha Values Included

Variables	EI	GI	ESC	GSC	NA	MI
EI	-	.20	.52***	.11	-.04	.12
GI	.28***	-	.35	.70***	.14	.02
ESC	.39***	.42***	-	.62***	.14	.00
GSC	.16**	.73***	.76***	-	.17	-.01
NA	-.01	.01	.23***	.16**	-	-.38***
MI	.06	.10*	.05	.04	-.14**	-
Previous[a] alpha	.89	.81	.85	.88	.93	.95
Current alpha	.92	.81	.84	.87	.93	.94

Note. Correlations among the six factors as reported by Picho & Brown (2011) in their initial development of the SIAS are shown above the diagonal, and correlations for the six factors that emerged from the current study are presented below the diagonal. Cronbach's alpha values can be found in the last two rows.

[a]Alpha values found in Picho & Brown's factor analysis of the SIAS using their sample

*$p < .05$. **$p < .01$. ***$p < .001$.

Appendix B

Table 2

Previous and Current Analyses of the SIAS' Sample Characteristics as Percentages

Sample Demographics	Picho & Brown's First Sample (*N*=206)	Picho & Brown's Second Sample (*N*=200)	Current Study's Sample (*N*=516)
Race			
White	82.3	68.4	24.2
Black	2.7	13.4	21.5
Asian	6.5	8.0	20.0
Latino	4.3	3.2	22.5
Other	8.0	7.0	11.9
Gender			
Male	20.3	50.8	29.7
Female	79.0	49.2	70.3

Note. Under the race column, other refers to those that self-reported belonging to biracial, Native American, or other racial categories.

Appendix C

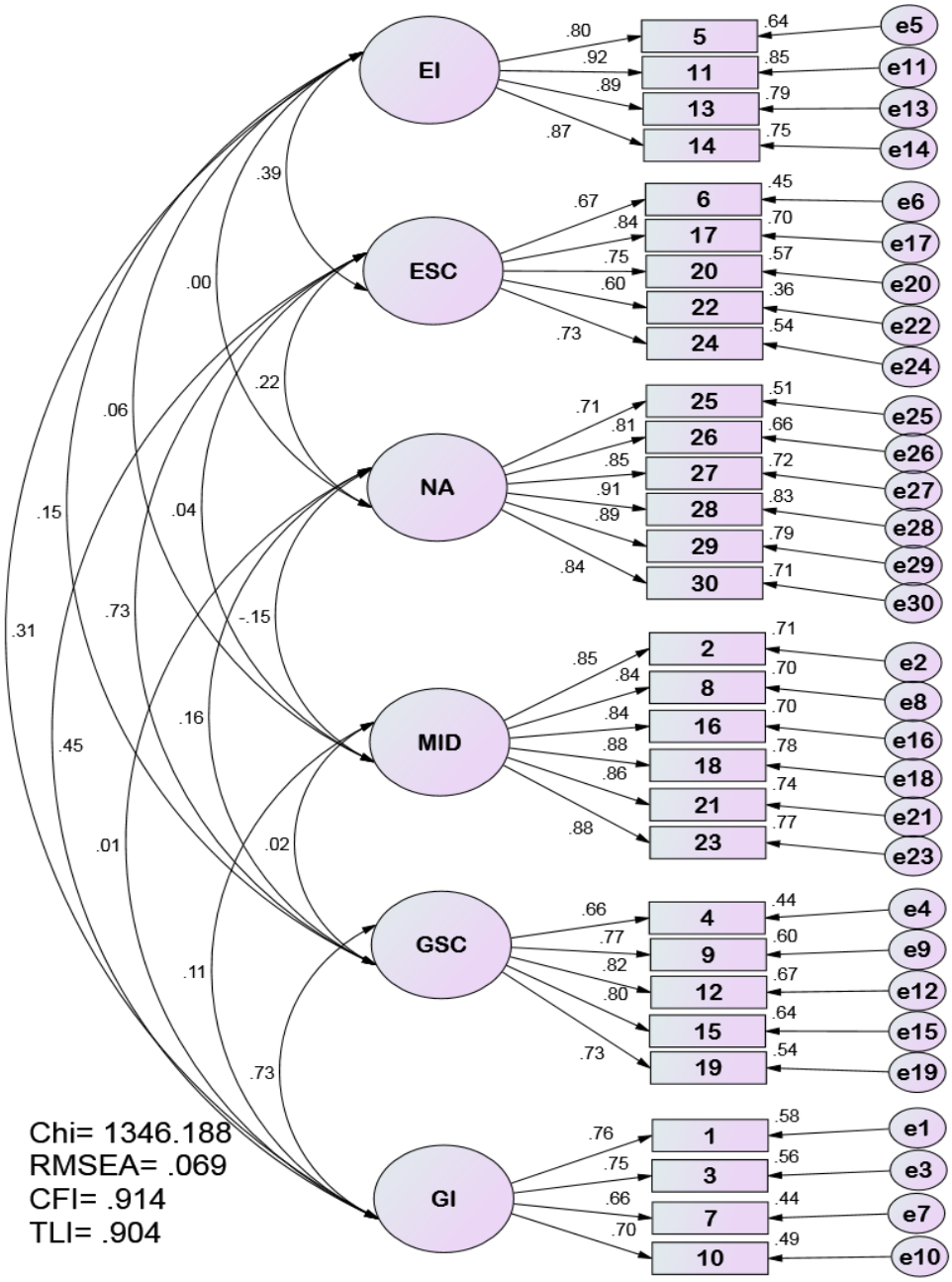

Figure 1. Results from the initial Confirmatory Factor Analysis on entire sample of 516 undergraduate students. This model does not include any freely estimated error terms.

Appendix D

Table 3

Fit Indices of Competing Models of the Social Identities and Attitudes Scale for Overall Sample

Model	χ^2	*df*	χ^2/df	$\Delta\chi^2$	Δdf	CFI	RMSEA
Initial	1346.19	390	3.45			.91	.069
Correlated Errors							
25 & 27	1225.47***	389	3.15	120.71	1	.93	.065
26 & 27	1158.57***	388	2.99	66.90	1	.93	.062
7 & 10	1092.76***	387	2.82	65.81	1	.94	.060
2 & 21	1032.49***	386	2.68	60.26	1	.94	.057
12 & 15	993.56***	385	2.58	38.93	1	.95	.055
25 & 26	927.00**	384		66.56	1	.95	.052

**$p < .01$. ***$p < .001$.

Appendix E

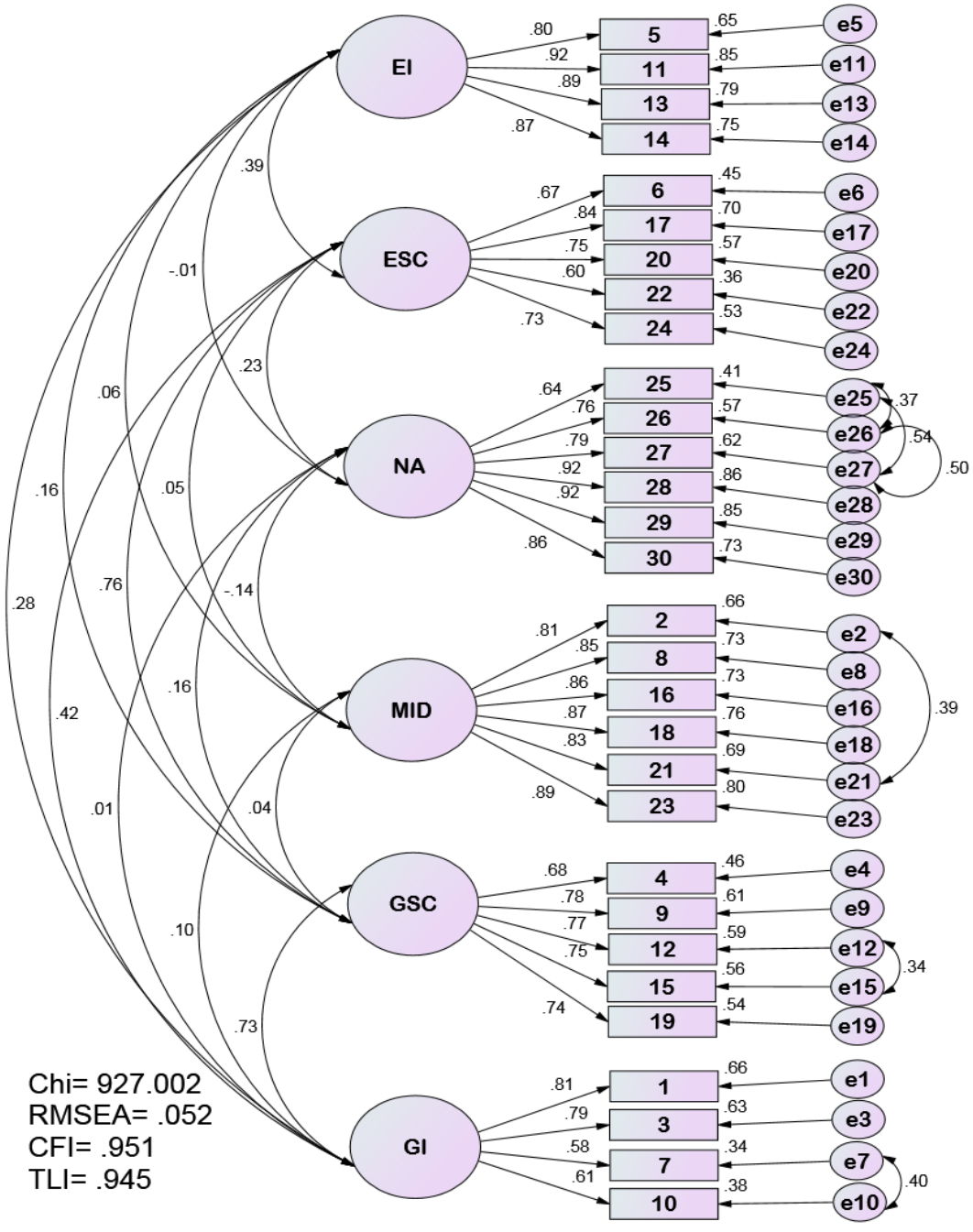

Figure 2. Results of the Confirmatory Factor Analysis after a total of six error correlations were freely estimated.

References

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. Journal of Experimental Social Psychology, 35(1), 29–46. doi:10.1006/jesp.1998.1371

Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51(6), 1173-1182.

Betancourt, H., & López, S. R. (1993). The study of culture, ethnicity, and race in American psychology. American Psychologist, 48(6), 629–637. doi:10.1037/0003-066X.48.6.629

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. Journal of Experimental Social Psychology, 39(6), 626–633. doi:10.1016/S0022-1031(03)00039-8

Byrne, B. M. (2013). Structural Equation Modeling With AMOS : Basic Concepts, Applications, and Programming, Second Edition (2nd ed.). Hoboken: Taylor and Francis.

Cadinu, M., Maass, A., Lombardo, M., & Frigerio, S. (2006). Stereotype threat: The moderating role of Locus of Control beliefs. European Journal of Social Psychology, 36(2), 183-197.

Croizet, J.-C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. Personality and Social Psychology Bulletin, 24(6), 588–594. doi:10.1177/0146167298246003

Davis, C., Aronson, J., & Salinas, M. (2006). Shades of threat: Racial identity as a moderator of stereotype threat. Journal of Black Psychology, 32(4), 399–417. doi:10.1177/0095798406292464

Feltz, D. L., Schneider, R., Hwang, S., & Skogsberg, N. J. (2013). Predictors of collegiate student-athletes' susceptibility to stereotype threat. Journal of College Student Development, 54(2), 184–201. doi:10.1353/csd.2013.0014

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. doi:10.1177/0146167202288010

Horn, J. L., & Mcardle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. Experimental Aging Research, 18(3), 117–144. doi:10.1080/03610739208253916

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3(4), 424–453. doi:10.1037/1082-989X.3.4.424

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. doi:10.1080/10705519909540118

Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. Journal of Educational Psychology, 99(3), 545–560. doi:10.1037/0022-0663.99.3.545

Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision making, and attention. Journal of Personality and Social Psychology, 99(3), 467–481. doi:10.1037/a0018951

Jameson, M., Diehl, R., & Danson, H. (2007). Stereotype threat impacts college athletes' academic performance. Current Research in Social Psychology, 12(5), 68-79.

Jordan, A. H., & Lovett, B. J. (2007). Stereotype threat and test performance: A primer for school psychologists. Journal of School Psychology, 45(1), 45–59. doi:10.1016/j.jsp.2006.09.003

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? Journal of Personality and Social Psychology, 83(3), 693–710. doi:10.1037/0022-3514.83.3.693

Kline, R. B. (2010). Principles and Practice of Structural Equation Modeling, Third Edition (3rd ed.). New York: Guilford Publications.

Lawrence, J. S., Marks, B. T., & Jackson, J. S. (2010). Domain identification predicts black students' underperformance on moderately-difficult tests. Motivation & Emotion, 34(2), 105–109. doi:10.1007/s11031-010-9159-8

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1(2), 130–149. doi:10.1037/1082-989X.1.2.130

Marx, D.M., & Stapel, D.A. (2006). Retracted: It's all in the timing: measuring emotional reactions to stereotype threat before and after taking a test. European Journal of Social Psychology, 36(5), 687-698.

Mazerolle, M., Régner, I., Morisset, P., Rigalleau, F., & Huguet, P. (2012). Stereotype threat strengthens automatic recall and undermines controlled processes in older adults. Psychological Science, 23(7), 723–727. doi:10.1177/0956797612437607

Nadler, J. T., & Clark, M. H. (2011). Stereotype threat: A meta-analysis comparing African Americans to Hispanic Americans. Journal of Applied Social Psychology, 41(4), 872–890. doi:10.1111/j.1559-1816.2011.00739.x

Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. Journal of Applied Psychology, 93(6), 1314–1334.

Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. Psychological Assessment, 7(3), 367–375. doi:10.1037/1040-3590.7.3.367

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. Journal of Consumer Research, 21(2), 381–391.

Picho, K., & Brown, S. W. (2011). Can stereotype threat be measured? A validation of the Social Identities and Attitudes Scale (SIAS). Journal of Advanced Academics, 22(3), 374–411. doi:10.1177/1932202X1102200302

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. Journal of Experimental Social Psychology, 38(2), 194–201. doi:10.1006/jesp.2001.1500

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. Journal of Personality and Social Psychology, 85(3), 440–452. doi:10.1037/0022-3514.85.3.440

Shapiro, J. R. (2011). Different groups, different threats: A multi-threat approach to the experience of stereotype threats. Personality and Social Psychology Bulletin, 37(4), 464–480. doi:10.1177/0146167211398140

Shapiro, J. R., Williams, A. M., & Hambarchyan, M. (2013). Are all interventions created equal? A multi-threat approach to tailoring stereotype threat interventions. Journal of Personality and Social Psychology, 104(2), 277–288. doi:10.1037/a0030461

Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. Social Justice Research, 20(4), 418–432. doi:10.1007/s11211-007-0047-7

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. doi:10.1006/jesp.1998.1373

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist, 52(6), 613–629. doi:10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. Journal of Personality and Social Psychology, 69(5), 797–811. doi:10.1037/0022-3514.69.5.797

Steele, J., James, J.B., & Barnett, R.C. (2002). Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. Psychology of Women Quarterly, 26(1), 46-50.

Steinberg, J. R., Okun, M. A., & Aiken, L. S. (2012). Calculus GPA and math identification as moderators of stereotype threat in highly persistent women. Basic & Applied Social Psychology, 34(6), 534–543. doi:10.1080/01973533.2012.727319

Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. Sociological Methodology, 8, 84–136. doi:10.2307/270754

Woodcock, A., Hernandez, P. R., Estrada, M., & Schultz, P. W. (2012). The consequences of chronic stereotype threat: Domain disidentification and abandonment. Journal of Personality & Social Psychology, 103(4), 635–646. doi:10.1037/a0029120