The Effects of Gender Stereotypic and Counter-Stereotypic Textbook Images on Science Performance

JESSICA J. GOOD
Rutgers University

JULIE A. WOODZICKA
Washington & Lee University

LYLAN C. WINGFIELD
Family Service of the Piedmont, Greensboro, North Carolina

ABSTRACT. We investigated the effect of gender stereotypic and counter-stereotypic images on male and female high school students’ science comprehension and anxiety. We predicted stereotypic images to induce stereotype threat in females and impair science performance. Counter-stereotypic images were predicted to alleviate threat and enhance female performance. Students read one of three chemistry lessons, each containing the same text, with photograph content varied according to stereotype condition. Participants then completed a comprehension test and anxiety measure. Results indicate that female students had higher comprehension after viewing counter-stereotypic images (female scientists) than after viewing stereotypic images (male scientists). Male students had higher comprehension after viewing stereotypic images than after viewing counter-stereotypic images. Implications for alleviating the gender gap in science achievement are discussed.

Keywords: education, gender differences, stereotypes, science

RESEARCH HAS DEMONSTRATED a distinct gender gap in the achievement scores in math and science of male and female students, as well as in interest in related career fields (Kahle & Lakes, 1983; Keller & Dauenheimer, 2003; Potter & Rosser, 1992; Turner & Bowen, 1999; U.S. Department of Labor, 2002). Beginning in high school, female students consistently under-perform compared
to male students on math and science portions of nearly all standardized tests (Shepardson & Pizzini, 1992; Sommers, 1996). Their lower scores on these standardized tests place female students at a disadvantage when applying to institutions of higher education. Once in college, female students enroll in fewer math and science courses and express less interest in pursuing math- and science-related careers than male students (Kahle & Lakes, 1983; Potter & Rosser, 1992; Turner & Bowen, 1999). Eleven out of the top 15 majors with the highest starting salaries in 2005 involved mathematics and science technology (Hargreaves, 2005), putting women at a disadvantage for future earnings due to their decreased interest in these fields.

Notably, this gender gap often does not emerge until high school (Marx & Roman, 2002), leading some to conclude that these differences in ability and preference between male and female students are not sex differences caused by genetic, chemical, or biological factors, but instead gender differences learned through socialization (Bleeker & Jacobs, 2004). Children learn these role expectations that may lead to the gender gap not only from their parents but from various forms of media, such as television cartoons and educational textbooks (Brownlow & Durham, 1997). Textbook images may depict a “hidden curriculum” of what is considered a perfect society, and this may have an effect on children’s academic and career interests (Bazler & Simonis, 2006; Potter & Rosser, 1992; Powell & Garcia, 1985; Shepardson & Pizzini, 1992). Ironically, although textbooks are considered vehicles for learning, they may actually hinder success in half of the student population.

The present study explores the impact of the subtle messages conveyed in textbooks through gendered pictures on the performance of female students in science. Specifically, we examined the influence of gender stereotypic textbook images as a potential medium for stereotype threat on students’ comprehension of a science lesson. We also examined the possibility of alleviating stereotype threat through the viewing of counter-stereotypic images.

The Gender Gap in Science and Math

Researchers have found a distinct gender gap in the math and science achievement scores of male and female high school students (Keller & Dauenheimer, 2003). The National Assessment of Educational Progress (NAEP) is the United States’ sole ongoing “indicator of what American students know and can do in major academic subjects” (National Center for Education Statistics, 2001, p. 27). Comparing scores on the NAEP across three decades, researchers found that although the gender gap in science scores has decreased since 1969, in 1999, males still outperformed females on the assessment at ages 13 and 17, but not at age 9 (Campbell, Hombo, & Mazzeo, 2001). Additionally, the gender gap in average science scores on the NAEP widened by 3 points at grade four and 5 points at grade eight between 1996 and 2000. Across several national standardized
tests, this gender difference in science scores has been shown to be small to moderate (Cohen’s $d = .32$), and consistent over the past 30 years (Hedges & Nowell, 1995). Additionally, a gender gap in interest in math and science develops around the time when students are preparing to enter high school (Marx & Roman, 2002). Researchers have found that between the ages of 9 and 13, a crucial change occurs in the attitudes of female students toward science. Girls at age 9 still desire active participation in scientific observations, but by age 13, they report a significantly decreased desire for such experiences (Kahle & Lakes, 1983; Potter & Rosser, 1992). By high school age, male students express greater interest in math and science fields than their female counterparts by enrolling in more math and science courses (Holden, 1987; Kahle & Lakes, 1983; Potter & Rosser, 1992; Smith, 1992). These findings support the idea that the gender gap in math and science does not fully materialize until upper-level grades, indicating that male and female students may be “taught” which academic pursuits are appropriate for them as they grow older (Bleeker & Jacobs, 2004; Campbell et al., 2001; Marx & Roman, 2002).

What are the implications of female students’ decreased achievement and interest in math and science? Lefevre, Kulak, and Heymans (1992) found that men are 4.5 times more likely to pursue college majors that are high in mathematical content. This is important because the top-paying majors for college graduates in 2005 were engineering, computer science, accounting, and information sciences (Hargreaves, 2005). The National Science Foundation (2000) reported that women accounted for only 23% of science and engineering occupations, with only 9% of employed engineers and 10% of employed physicists being women. While there has been evidence that the gender gap in math and science has narrowed since the 1980s, researchers are still finding that women generally have lower levels of confidence in their abilities in science-related areas than men (Bleeker & Jacobs, 2004). The current study investigated the possibility that textbook images are one way through which the gender gap in math and science is maintained.

**Images in Textbooks**

Textbooks have been indicted for perpetuating gender stereotypes through the images and language they use. A review of U.S. History textbooks found that in one textbook, there are four men for every woman mentioned, and less than 3% of the history is about women (Sadker & Sadker, 1994). In a review of science textbooks, researchers found that the texts contained minimal information regarding the accomplishments of women in science or about scientific topics of interest to women such as menstruation, childbirth, and menopause (Potter & Rosser, 1992). Women are not only underrepresented in text but also in images. Men are pictured more frequently in textbooks, while women remain relatively hard to find in illustrations, examples, and subject and author indices (Gray, 1977; Hogben & Waterman, 1997; Potter & Rosser, 1992). In an analysis of 80 science
textbooks from both elementary and secondary schools, men were pictured more often than women, and over 85% of the occupations pictured were embodied by men (Powell & Garcia, 1985). Most recently, in a review of seven high school chemistry textbooks, researchers found that only one book achieved gender parity in image representation; the other books overwhelmingly pictured more men than women (Bazler & Simonis, 2006).

Beyond basic frequency of representation, researchers have also examined both the explicit and implicit derogatory ways in which women are portrayed when they are pictured. Hogben and Waterman (1997) found that men were more often portrayed as dominant or having higher status than women in textbook images. For instance, when a doctor and nurse are pictured together, the doctor is typically shown as a man, while the secondary role of the nurse is typically shown as a woman. Women are also more likely to be shown as passive actors merely reacting to their environment, while men are shown as active in altering the situation around them. In their assessment of Introductory Psychology and Human Development textbooks, Peterson and Kroner (1992) found that men were significantly more likely to be shown in active roles by directly engaging in and initiating activity (e.g., leading a group). In this way, men are depicted as being active agents, managing and controlling their environment. Conversely, women were more likely to be depicted as passive, reactive to their environment, and the recipients of others’ action (e.g., watching others).

**Stereotype Threat**

Exactly how do stereotypic images in textbooks affect female perceptions of and performance on math- and science-related tasks? Stereotype threat may be one potential mechanism for reducing women’s performance and interest in math and science areas. Stereotype threat is a well documented phenomenon by which individuals, fearful of confirming a negative stereotype about their group, display decreased performance on a task relevant to the negative stereotype (Steele & Aronson, 1995). For example, Steele and Aronson (1995) found that Black students performed worse on a test supposedly “diagnostic” of their intellectual ability than on a “non-diagnostic” test, even though the same test was used in both conditions. Of particular relevance to the present study, O’Brien and Crandall (2003) found that women given a math test that had “been shown to produce gender differences” scored worse than women taking the same test that had “not been shown to produce gender differences” (O’Brien & Crandall, 2003, p. 785). By priming the female participants with the stereotype that men are innately superior at math and science related tasks the experimenters induced a condition of stereotype threat, causing the women to perform more poorly on the subsequent difficult math test. The phenomenon of stereotype in regard to the stereotype of women’s math ability has been well documented (Brown & Josephs, 1999; Smith & White, 2002; Spencer, Steele, & Quinn, 1999).
Importantly, the effect of stereotype threat on women has also been replicated in non-laboratory settings, lending ecological validity to this concept. Keller and Dauenheimer (2003) found that female high school students taking a math test in mixed-gender groups within their own classroom-setting scored worse when the math test was said to show gender differences than when it was said to not show gender differences. The finding that stereotype threat can be induced in natural settings lends support to the possibility that female students experience threat as a daily part of their academic career (Keller & Dauenheimer, 2003; Smith & White, 2002).

Although past research has generally not applied stereotype threat to the realm of images in text, the role of advertisements in producing stereotype threat and subsequent performance deficits has been examined (Davies, Spencer, Quinn, & Gerhardstein, 2002; Davies, Spencer, & Steele, 2005). Women who viewed stereotypic television commercials not only scored lower on a math test than women who viewed neutral ads, but also were more likely to avoid math questions in favor of verbal questions. Importantly, women who viewed counter-stereotypic commercials scored as well as the men in the same condition. If viewing television commercials is sufficient to induce stereotype threat and avoidance of traditionally masculine domains in female students, it is possible that viewing stereotypic pictures in textbooks is also sufficient to induce stereotype threat and contribute to impaired performance in math- and science-related fields.

An important caveat to note when considering stereotype threat is that the absence of stereotype threat does not completely eliminate the score-gap between groups. In many of the studies presented above and in Steele and Aronson’s (1995) seminal article, the gap in scores between men and women (or Black and White students) was still present in the no-threat conditions; however, the magnitude of this gap was significantly lower than in the stereotype threat conditions (Sackett, Hardison, & Cullen, 2004). Thus, we do not propose that eliminating biased textbook images and thereby alleviating stereotype threat will completely erase the gender gap in math and science scores, but we do think the gap will be lessened.

When examining possible interventions aimed at raising female students’ scores in math and science, it is important to take into consideration the effects of these actions upon male students as well. Marx and Roman (2002) found that male students scored lower on a math test when administered by a female experimenter believed to be highly competent in math than when administered by a male experimenter or a female experimenter perceived not to be competent in math. It is possible that seeing a counter-stereotypic female role model, such as in a textual illustration or photograph, could impair male students’ test performance. The present study includes both female and male students in an attempt to determine the degree to which stereotypic and counter-stereotypic images impact both sexes.
Pilot Study

A pilot study was conducted in order to test the manipulation of images within a text-based science lesson. Several science lessons were created, with consistent text but varying stereotypic image content. After reading the lesson, participants were tested for comprehension of the lesson. Effects of the image manipulation were marginally significant with an undergraduate sample; data from the stereotypic and counter-stereotypic conditions revealed that male performance was enhanced by stereotypic images, and female performance was not significantly impacted by image content. Manipulation checks revealed that photo content may have lacked clarity, and thus image quality and content were improved for the present experiment. Additionally, the college sample may have been too old for the manipulation to have an effect, due to possible disidentification from science as a subject. In order for stereotype threat to have the greatest effects, group members should be highly identified with the task (Steele & Aronson, 1995). Cokley (2002) found that self-esteem became increasingly detached from the academic performance of African Americans, a stereotyped group, the longer they stayed in school. College age students may be too advanced in their academic careers, many already having chosen their career paths. To try and capture a population that may still be identified with science, we used a younger sample in the current research, specifically ninth and tenth grade students. However, because the gender gap in math and science may begin between the ages of 9 and 13 (Holden, 1987; Marx & Roman, 2002; Smith, 1992), we also included a measure of students’ interest in science and scientific careers in order to control for disidentification.

The Present Research

This research examines the effects of stereotypic and counter-stereotypic textbook images on high school students’ science comprehension and state anxiety. We expected that female students exposed to counter-stereotypic images would perform better on the comprehension assessment and report less anxiety than female students exposed to stereotypic images. The opposite was predicted for males; male students who viewed stereotypic images were expected to have higher comprehension and less state anxiety than male students viewing counter-stereotypic images. We also expected that female and male students exposed to mixed gender images would perform more equivalently on the comprehension assessment and report experiencing comparable levels of anxiety.

Method

Participants

The participants were 81 9th- and 10th-grade students (29 male students, 52 female students) who had not taken a chemistry course nor had specific instruction
in chemical reactions. Age ranged from 13–17 years ($M = 14.81$ years), and 97.5% self-identified as Caucasian. Participants were enrolled in either honors ($n = 36$) or academic level ($n = 45$) classes and were recruited through letters sent home to their parents by the school principal. Due to the age of the participants, we felt it necessary to control for disidentification with the scientific domain. Therefore, three participants were eliminated due to incomplete data regarding interest in science.

**Stimulus Materials**

Three versions of a chemistry lesson on chemical reactions and equations taken from a high school chemistry book were used (Dingrado, Gregg, Hainen, & Wistrom, 2002). Each excerpt consisted of three pages of textual information and three manipulated photographs. The photographs varied based upon condition, while the textual information remained constant across conditions. In the counter-stereotypic condition, photos consisted of three lone female scientists. In the stereotypic condition, images consisted of three lone male scientists. In the mixed gender condition, photos consisted of one lone female scientist, one lone male scientist, and one image of both a male and a female scientist working together in equally active roles. Photographs were obtained from online databases. Scientists in the photographs were identifiable as such by their lab coats and use of chemistry equipment, latex gloves, safety glasses, etc. In order to hold non-gender factors constant across all conditions, all persons pictured were Caucasian, approximately middle-aged, and moderately attractive.

**Dependent Variables**

*Comprehension assessment.* A comprehension test consisting of 12 multiple choice and fill-in-the-blank questions was created by the researchers to assess participants’ comprehension and retention of the textbook lesson. Sample items included “The starting substances of a chemical reaction are called ____” and “Which of the following is not evidence that a chemical reaction has taken place?” Internal scale reliability was acceptable, $\alpha = .76$. The test assessed all major points covered in the lesson; it did not assess any information that was not present in the lesson.

*State anxiety scale.* The 20-item Spielberger State Anxiety Inventory (Spielberger, 1985) was used to assess participants’ level of anxiety following the testing situation. Participants were asked to rate on a 4-point scale (1 = not at all, 4 = very much so) how they were feeling at the moment they were answering the questionnaire (e.g., “I feel calm”). Internal scale consistency was high, $\alpha = .89$. The Spielberger State Anxiety Inventory has been shown to be a valid measure of in-the-moment anxiety (Novy, Nelson, Goodwin, & Rowzee, 1993).
Academic and career survey. A survey was created by the researchers to investigate the current academic and future career interests of the participants. Participants were asked to rate on a 5-point scale (1 = not at all, 5 = extremely) the degree to which they liked several academic subjects and the likelihood that they would pursue a career related to those academic fields (e.g., English, math, science, art). The primary measure of concern was interest in science; however, other fields of study were included to obscure the focus on science. The survey was administered to students both prior to and following the lesson and comprehension test.

Manipulation check. A multiple choice question was designed to assess participants’ ability to identify the images in the text. Participants were specifically asked to identify the content of the first picture shown and were given four response options (i.e., male scientist, female scientist, male and female scientists, chemistry equipment). Questions were also included as an attempt to determine the level of difficulty of the lesson and comprehension test as perceived by the participants as well as the amount of effort participants exerted.

Procedure

The study was administered to students by their regular 9th- and 10th-grade science teachers during one class period. Teachers passed out manila envelopes to each participating student. The envelopes contained all of the testing materials. Students were randomly assigned to one of the three conditions based on which envelope they received. After reading an instruction script, the teachers asked for any questions from the students and then directed the students to open the envelopes and begin the study.

Upon opening the envelopes, students removed two stapled packets. The first packet contained a page of directions, the Academic and Career Survey, and one of the three chemistry lessons (stereotypic, counter-stereotypic, or mixed gender). The last page of the packet instructed students to place the packet face down on their desk and then begin the second packet. The second packet began with the Comprehension Assessment. Following this test, students completed the State Anxiety Inventory, a second Academic and Career Survey, and the Manipulation Check. The last page of the second packet instructed students to place their completed testing materials back into the envelope and seal it. Following completion of the study, teachers collected the testing materials and returned them to the experimenters.

Results

Manipulation Check

Across the three conditions, 84% of the participants correctly recalled the content of the images they viewed. A 2 (sex) × 3 (image condition) ANOVA
revealed no significant main effects or interactions among the groups in their ability to correctly identify image content. Independent t-tests revealed that participants who recalled the textbook images correctly did not differ significantly on any of the dependent measures from those who did not recall the images correctly. Further, results from analyses of only the participants who answered the manipulation check question correctly did not differ from results obtained from the full data set. Because there were no differences, we used the full data set for all subsequent analyses.

A question to determine amount of effort put forth by students on the comprehension test was also included. Students were asked to rate how hard they tried on a 7-point scale, with 1 being not at all and 7 being tried my best. A 2 (sex) × 3 (image condition) ANOVA on effort revealed no significant differences between groups (M = 5.35, SD = 1.61).

Comprehension Assessment

Means were computed for male and female students in each condition based on the number of correct answers out of 12 total questions (see Table 1 for means, standard deviations, and number of participants in each condition). Across conditions, comprehension scores ranged from 0 to 12 (M = 8.24, SD = 2.76). To determine any differences in comprehension scores between conditions, a 2 (sex) × 3 (image condition) between subjects ANCOVA was computed. Because of the possible confounding effect of disidentification with science, initial

| TABLE 1. Mean Comprehension Scores According to Sex and Image Condition |
|-----------------|---|---|---|
|                 | Female | Male | Total |
| **Image Condition** |         |     |     |
| Stereotypic      | 7.42 (SD = 3.35) | 9.00 (SD = 2.18) | 7.86 (SD = 3.11) |
|                  | n = 18 | n = 7 | n = 25 |
| Counter-Stereotypic | 9.38 (SD = 1.88) | 7.70 (SD = 1.72) | 8.73 (SD = 1.97) |
|                  | n = 16 | n = 10 | n = 26 |
| Mixed Gender     | 8.37 (SD = 3.30) | 8.25 (SD = 3.20) | 8.31 (SD = 3.19) |
|                  | n = 15 | n = 12 | n = 27 |
| **Total**        | 8.35 (SD = 2.99) | 8.24 (SD = 2.50) | 8.31 (SD = 2.80) |
|                  | n = 49 | n = 29 | n = 78 |

*Note. Comprehension scores are out of possible high score of 12.*
interest in science (measured before the lesson) was covaried. A significant interaction of condition and sex was found, $F(2, 71) = 3.78, MS = 23.98, p = .027$. As predicted, female students scored significantly higher in the counter-stereotypic condition ($M = 9.38, SD = 1.88$) than in the stereotypic condition ($M = 7.42, SD = 3.35$), $t(32) = 2.07, p = .047, d = .72$. Male students, on the other hand, demonstrated higher comprehension in the stereotypic condition ($M = 9.00, SD = 2.18$) than in the counter-stereotypic condition ($M = 7.70, SD = 1.72$). Although this difference was not statistically significant ($p = .19$), the effect size was large ($d = .66$).

It was also hypothesized that female and male students in the mixed condition would perform equally on the comprehension task. As expected, an independent $t$-test revealed no significant differences between the comprehension scores of male and female students in the mixed condition ($p = .927, d = .04$). See Figure 1 for a graphical representation of the results.

**State Anxiety**

It was hypothesized that female students viewing counter-stereotypic images would report less anxiety than female students exposed to stereotypic images; in contrast, male students exposed to stereotypic images would report less anxiety than male students who viewed counter-stereotypic images. A $2 \times 3$ (image condition) factorial ANOVA revealed a significant interaction of condition and sex, $F(2, 71) = 3.78, MS = 23.98, p = .027$. As predicted, female students scored significantly higher in the counter-stereotypic condition ($M = 9.38, SD = 1.88$) than in the stereotypic condition ($M = 7.42, SD = 3.35$), $t(32) = 2.07, p = .047, d = .72$. Male students, on the other hand, demonstrated higher comprehension in the stereotypic condition ($M = 9.00, SD = 2.18$) than in the counter-stereotypic condition ($M = 7.70, SD = 1.72$). Although this difference was not statistically significant ($p = .19$), the effect size was large ($d = .66$).

It was also hypothesized that female and male students in the mixed condition would perform equally on the comprehension task. As expected, an independent $t$-test revealed no significant differences between the comprehension scores of male and female students in the mixed condition ($p = .927, d = .04$). See Figure 1 for a graphical representation of the results.

**State Anxiety**

It was hypothesized that female students viewing counter-stereotypic images would report less anxiety than female students exposed to stereotypic images; in contrast, male students exposed to stereotypic images would report less anxiety than male students who viewed counter-stereotypic images. A $2 \times 3$ (image condition) factorial ANOVA revealed a significant interaction of condition and sex, $F(2, 71) = 3.78, MS = 23.98, p = .027$. As predicted, female students scored significantly higher in the counter-stereotypic condition ($M = 9.38, SD = 1.88$) than in the stereotypic condition ($M = 7.42, SD = 3.35$), $t(32) = 2.07, p = .047, d = .72$. Male students, on the other hand, demonstrated higher comprehension in the stereotypic condition ($M = 9.00, SD = 2.18$) than in the counter-stereotypic condition ($M = 7.70, SD = 1.72$). Although this difference was not statistically significant ($p = .19$), the effect size was large ($d = .66$).

It was also hypothesized that female and male students in the mixed condition would perform equally on the comprehension task. As expected, an independent $t$-test revealed no significant differences between the comprehension scores of male and female students in the mixed condition ($p = .927, d = .04$). See Figure 1 for a graphical representation of the results.
condition) between subjects ANCOVA, covarying initial science interest, failed to reveal significant main effects or a significant interaction \((p > .05)\). It was also hypothesized that male and female students in the mixed condition would report similar levels of anxiety. As expected, an independent \(t\)-test found no significant differences between the levels of anxiety reported by males and females in the mixed condition \((p = .21, d = .52)\). The moderate effect size however, indicates that female students did report greater anxiety than male students in the mixed-gender condition.

We tested whether state anxiety mediated the relationship between textbook image condition and comprehension test scores (see Baron & Kenny, 1986). Results of the three regressions (anxiety on textbook image condition, comprehension score on image condition, and comprehension score on image condition and anxiety) indicate that state anxiety did not act as a mediator.

**Academic and Career Interests**

Academic interest in science was measured in order to control for students’ disidentification with the scientific domain. Post measures of interest were also collected to determine any changes in academic and career interest following the experimental manipulation. A 2 (sex) \(\times\) 2 (pre-post) \(\times\) 3 (image condition) mixed group ANOVA with sex and image condition as the between subjects variables and academic interest (pre and post-test) as the within subjects variable was computed. A significant main effect of time was found, \(F(1, 72) = 8.36, MS = 1.17, p = .005, d = .19\), with students reporting greater interest prior \((M = 3.62, SD = 1.06)\) to the lesson than afterwards \((M = 3.42, SD = 1.09)\). Although students reported decreased academic interest in science after reading the lesson and taking the subsequent test, a 2 (sex) \(\times\) 2 (pre-post) \(\times\) 3 (condition) mixed group ANOVA revealed no significant main effects or interactions for their desire to pursue a career in science \((p > .05)\). To clarify the findings regarding decreased interest in science following the experiment, analyses were run on each academic subject listed on the survey. Students also reported decreased interest in English \((pre-M = 3.11, SD = 1.06, post-M = 2.92, SD = 1.22, d = .17)\) and math \((pre-M = 2.81, SD = 1.16, post-M = 2.69, SD = 1.17, d = .10)\) following the lesson and comprehension test, \(ps < .05\); these effects were not found for music, social studies, and art.

**Discussion**

The present study investigated the effect of stereotypic and counter-stereotypic textbook images on high school students’ comprehension of a difficult science lesson. A mixed gender condition was also tested in an attempt to elucidate a form of image presentation which would equalize performance across genders. As predicted, female students viewing counter-stereotypic images were found to have significantly higher comprehension of the science lesson.
than female students viewing stereotypic images. The means were in the opposite direction for male students; although not statistically significant, a large effect size was found. When viewing mixed gender images, no difference was found in the comprehension of male and female students. Female students reported more anxiety than male students in the mixed gender condition, but this difference was not statistically significant. Students reported a decreased interest in science as an academic subject, as well as math and English, following exposure to the lesson and comprehension test, but did not alter their interest in pursuing a career in science.

Research has shown that when a positive stereotype about an ingroup or a negative stereotype about an outgroup is made salient, ingroup members may experience stereotype lift, and subsequently perform better on evaluative measures (Shih, Pittinsky, & Ambady, 1999; Walton & Cohen, 2003). Male students, after incorporating gender into their task of reading the science lesson, may have been reminded of the stereotype that men are superior at math and science, and thus scored higher in the stereotypic condition than in the counter-stereotypic or mixed gender conditions. In contrast, female students, after viewing images counter to the negative stereotype about the ingroup, demonstrated greater comprehension than after viewing traditionally stereotyped images. The mixed gender condition, which contained pictures of single male and female scientists, as well as a picture of a male and female scientist working together, was sufficient to equalize comprehension scores across the sexes. In addition, comprehension scores for both male and female students in the mixed gender condition fell between scores in the stereotypic and counter-stereotypic conditions. This finding is especially important, as it provides evidence that the mixed gender condition did not simply represent the absence of threat, in which case we would expect men to still score slightly higher than women (Sackett, Hardison, & Cullen, 2004). Instead, the effect size indicates that there was very little difference between the scores of male and female students. Although one must keep in mind the between-subjects design of the experiment, the mixed gender condition did appear to equalize performance between male and female students. Providing students with diverse role models within textbook images may ensure that no one group of students is affected by stereotype threat. Future research should more fully examine the impact of diverse role models as a way to reduce the negative effects of textbook images on female students, as well as other negatively stereotyped groups.

Since Steele and Aronson (1995) originally defined stereotype threat, anxiety has been considered the main mediator for the phenomenon. However, in the current research, anxiety did not act as a mediator. This finding is consistent with results obtained by other researchers who have not found significant effects of anxiety under stereotype threat conditions (Bosson, Haymovitz, & Pinel, 2004; Keller & Dauenheimer, 2003). Unfortunately, we did not include in our study a baseline measure of anxiety with which to compare post-test anxiety among
students. It is possible that the difference in anxiety may mediate the effects of stereotype threat. Future research should investigate this possibility.

Another possible mediator that may have impacted comprehension scores in the current research is cognitive load. For example, Schmader and Johns (2003) hypothesized that decreases in performance under conditions of stereotype threat may be due to reductions in working memory capacity. Noting that anxiety and stress can increase cognitive load and thereby reduce working memory capacity, Schmader and Johns (2003) found that women told they were taking a test of quantitative capacity that had been shown to produce gender differences recalled fewer words and were less accurate on a math test than women told they were taking a test of working memory capacity. Analyses indicated that working memory capacity mediated the effect of stereotype threat on math test performance in the above experiments. Thus, it is possible that in the current study female students in the stereotypic condition were unable to block certain thoughts about the negative stereotype, which led to an increased cognitive load and decreased working memory capacity, impairing their ability to remember the content of the science lesson. Unfortunately, cognitive load was neither measured nor manipulated in the present research, thus speculations about the role of cognitive load as a mediator are necessarily tentative.

It is also possible that the process was less cognitive and more automatic or passive. Cialdini & Goldstein (2004) reviewed the recent literature on compliance and conformity and concluded that people may passively conform to societal ideals or stereotypes in an attempt to maintain their self-concept or feel that they fit in with a group. In our study, activating the stereotype that women are bad at math and science may have led women to passively conform to that view in order to fit with societal norms. Future research should investigate this type of conformity as a less cognitive mediator of stereotype threat.

Following the science lesson, participants across conditions reported decreased interest in science, math, and English, but not music, social studies, or art. The lesson involved science (chemistry) and math (adding chemical equations), as well as elements of English (reading comprehension). Engaging in this difficult activity may have led students to report decreased interest in these domains, and not in domains not associated with the lesson. Although no long term measure was used, we do not believe the lesson caused any lasting decrements in student interest.

Generally participants were not able to report perfect awareness of image content. However, awareness of image content or threat may not be necessary for the deleterious effects of stereotype threat to occur. It is precisely this unawareness that may make textbook images an insidious threat to students’ learning. Schmader and Johns (2003) concluded that a conscious awareness of threat may not be necessary for stereotype threat to have an effect on working memory capacity and cognitive performance. This suggests that textbook images can have an impact on students’ comprehension of the text, without the
students’ being consciously able to pinpoint a reason for their reduced learning. The danger then lies in the lifelong disadvantage girls face as a result of decreased learning in their childhood, causing many to credit girls with less academic ability than boys, and perpetuating the stereotype once more. Without the ability to consciously recall the gender-biased nature of textbook images, girls may begin to believe in the stereotype, attributing their inferior performance to a supposed natural sex difference, and ultimately leading them to avoid majors and professions involving math and science. Thus, textbooks, considered positive instruments of learning, may ironically teach girls that they have no place in the academic areas of math and science and thereby reduce their achievement and enjoyment within these disciplines. The present research taps only one moment in time; future research is needed to determine the cumulative effect of gender-biased textbooks throughout students’ academic career. It would also be useful to understand when biased images or textbook content have the most impact on students’ interest in science and beliefs about their scientific abilities.

We do not assume that gender-stereotypic textbook images are the sole cause of the gender gap in math and science achievement. Clearly other forces are at work in addition to textbooks. A limitation of the current research is the small sample size. However, the current study has shown that textbook images do have some impact on students’ retention of lesson material, and thus may play a role in maintaining the gender gap. The mixed gender image condition suggests a tentative solution to this problem. It is encouraging that male and female students viewing pictures of active male and female scientists working together demonstrated approximately equal comprehension of the lesson. Future research is necessary to further elucidate the effects of photographs in textbooks on students’ performance and interest in scientific subjects. Research on mixed-gender textbook images is especially needed, as they may represent a simple and cost-effective way to remedy the negative effects of stereotypic textbook images. Research should investigate the influence of diverse role models presented in textbooks as a way of improving performance of multiple stereotyped groups, not just women. Although eliminating gender bias in textbooks will most likely not eradicate the gender gap in science interest and achievement, it will begin to chip away at an ever crumbling foundation.

**AUTHOR NOTES**

Jessica J. Good is a member of the psychology department at Rutgers University, New Brunswick, New Jersey. Julie A. Woodzicka is a member of the psychology department at Washington & Lee University, Lexington, Virginia. Lylan C. Wingfield is the director of Substance Abuse Services at Family Service of the Piedmont, Greensboro, North Carolina.
REFERENCES


Good, Woodzicka, & Wingfield 147


*Received August 16, 2007*

*Accepted June 16, 2008*