CrossMark

Stereotype threat and gender differences in chemistry

Cijy Elizabeth Sunny[1] · Gita Taasoobshirazi[1] ·
Lauren Clark[2] · Gwen Marchand[3]

Abstract Stereotype threat theory (STT) offers one explanation for achievement differences in math and science for both women and minority students. Specifically, STT posits that the perceived risk of confirming a negative stereotype about an individual's identity group acts as a psychological burden that negatively impacts performance. This study examined the impact of stereotype threat (ST) on gender differences in chemistry achievement, self-efficacy, and test-anxiety using a four-group, quasi-experimental design. 153 introductory-level college chemistry students were randomly assigned to one of four ST conditions including an explicit ST condition, an implicit ST condition, a reverse ST condition, and a nullified condition. Results indicated that there were no gender differences by ST condition; however, overall, the men had higher self-efficacy and lower test-anxiety than the women. An analysis of open-ended questions asking students about their intent to major in chemistry, beliefs regarding barriers to their achievement on the chemistry test, and gender differences in opportunities and mental capacity to achieve in chemistry provided insight into the quantitative results. Implications of our findings for future research on ST are discussed.

Keywords Stereotype threat · Chemistry · Gender differences · Self-efficacy · Test-anxiety

There is a need for approximately one million more college graduates in Science, Technology, Engineering, and Mathematics (STEM) fields according to the President's Council

✉ Gita Taasoobshirazi
   gita.tshirazi@gmail.com

1   Department of Quantitative and Mixed Research Methodologies, University of Cincinnati, Cincinnati, OH, USA

2   Department of Developmental and Learning Sciences, University of Cincinnati, Cincinnati, OH, USA

3   Department of Educational Psychology & Higher Education, University of Nevada, Las Vegas, Las Vegas, NV, USA

🍎 Springer

of Advisors on Science and Technology (PCAST) report from Washington D.C., United States (Olson and Riordan 2012). Increasing the supply side of the STEM pipeline in order to increase the STEM work force has been a major challenge (Wang 2013). This pipeline is sometimes called leaky as it is often unable to retain students from secondary school all the way to STEM careers, and a large number of the students lost from the pipeline are women (Blickenstaff 2005; Chen and Soldner 2013).

Although more women are attending college and pursuing degrees in the sciences compared to previous decades, they are still underrepresented in many sciences and in more advanced science degrees (Hill et al. 2010; National Science Foundation (NSF) 2015). Numerous reasons have been reported in the research over the years to explain the gender disparity in science achievement and participation including differences in motivation, differences in self-efficacy, lack of role models for women, and differences in parent and teacher support (e.g., Desouza and Czemiak 2002; Enman and Lupart 2000; Greene and DeBacker 2004; Mattern and Schau 2002; She 2001; Shin and McGee 2002; Tenenbaum and Leaper 2003). Stereotype threat (ST), or when a situation poses a risk by which one's behavior could be interpreted as confirming a negative stereotype about a person's social group, has been hypothesized to contribute to gender differences in science achievement and participation (Marchand and Taasoobshirazi 2013; Smith 2004; Steele 1997). Although there is substantial research on the impact of ST on the performance of women in mathematics (Johns et al. 2005; O'Brien and Crandall 2003; Schmader 2002; Spencer et al. 1999), comparatively fewer studies have examined the role of ST on gender differences in science (Marchand and Taasoobshirazi 2012). Furthermore, there is a dearth of research exploring the impact of ST on gender differences in chemistry, particularly at the post-secondary level.

More than a simple expansion into a new domain, studying ST with respect to chemistry offers a contrast to past research because the gender-landscape within fields such as biology and chemistry are shifting more rapidly than in previously-studied fields such as physics or mathematics. For instance, 2008 represented the first year that women earned more doctorate degrees in biology than men (Matson 2013). The data on degrees awarded in chemistry also indicates a closing of the gender gap, especially for less advanced degrees. Although women only hold 39 % of the doctorates in chemistry, they hold nearly half of the bachelor's degrees awarded in chemistry (Matson 2013). Because ST researchers have posited that both role models and an increased sense of belonging will mitigate ST (Steele 1997), fields like chemistry or biology offer unique testing grounds to better understand the impact of ST on achievement. The purpose of this study was to offer such an extension of the research on gender differences in science and ST and to examine whether ST impacts differences in the performance, self-efficacy, and test-anxiety of women and men in chemistry.

Despite the closing of the gender gap in chemistry participation for bachelor degrees, there are still reports of gender differences in performance at the undergraduate level and in participation in more advanced chemistry degrees (Matson 2013; Stieff et al. 2012). If ST threat is a factor contributing to those differences, interventions can be implemented in undergraduate, introductory-level chemistry courses to minimize those negative effects.

Stereotype threat

The origins of stereotype threat research in education may be traced to the investigation of racial differences in performance in standardized testing situations (Steele and Aronson 1995). The notion that stereotypes held about a particular group may create psychologically threatening situations associated with fears of confirming judgment about one's

group, and in turn, inhibit learning and performance (Johnson et al. 2012), has since been extended to explore a variety of gender and racial group differences across domains such as athletics, chess, and mathematics (Schmader 2002; Smith 2004; Spencer et al. 1999).

Effects of ST on performance

Research on the impact of ST on females in science has been scarce (Steele 1997). Much of what we know about gender and ST comes from the research in mathematics. Extensive research has been conducted on the impact of ST on women's performance in mathematics. Several combinations of experimental or quasi-experimental conditions have been used to assess the impact of ST on gender differences in mathematics. ST has been studied in conditions where the test is described as diagnostic (e.g., you are taking a math test) or non-diagnostic (e.g., this is a problem solving task) (Johns et al. 2005; Kiefer and Seka-quaptewa 2007). ST has also been studied in conditions when the threat is made implicit (e.g., just being in an everyday mathematics testing situation), explicit (e.g., students are told men perform better than women on a test), or nullified (e.g., equating the groups) (Smith and White 2002). Most commonly, implicit or explicit ST conditions are compared with a nullified condition (e.g., O'Brien and Crandall 2003). For example, Spencer et al. (1999) compared college-level men and women's mathematics performance across two conditions. Students were told, prior to taking a mathematics test, that there were no gender differences on the test (a nullified condition) or were given no information regarding gender differences on the test (implicit ST condition). Results indicated that the men outperformed the women in the implicit ST condition, but that gender differences disappeared in the nullified ST condition. These results suggested that when no information was given, women still underperformed because of the existing and implicit stereotype that women are less capable than men in mathematics (O'Brien and Crandall 2003; Quinn and Spencer 2001) and that the testing situation alone was enough to trigger the threat.

Although studies in the domain of mathematics inform our understanding of gender and ST in STEM, it is not clear as to whether effects generalize to science content areas, such as chemistry. A detailed search of the research on ST and science/chemistry led to the retrieval of approximately three empirical studies examining the impact of ST on gender differences in science performance across ST conditions. Only one of those studies examined ST in chemistry. These three studies are described below.

The research on ST in chemistry has studied how exposing high school chemistry students to images of scientists impacts their performance (Good et al. 2010). Good et al. (2010) created three conditions where students were exposed to a chemistry text including images of all male scientists, all female scientists, or a mixed group of scientists. Results indicated that the women performed best on the chemistry test in the 'all female' image condition. The men performed best in the 'all male' image condition. The men and women performed similarly in the 'mixed gender' image condition. The authors also examined whether test-anxiety was impacted by the ST conditions, but found no significant effects for test-anxiety.

Bell et al. (2003) assigned college engineering students to three ST conditions including a diagnostic condition (the test is measuring engineering aptitude), non-diagnostic condition (test responses are being used to modify and improve the test), and a gender fair condition (test is a measure of aptitude, but men and women have been found to perform equally well on the test). Results indicated that the men outperformed the women in the

diagnostic condition but not the non-diagnostic and gender fair conditions. For the women, the instructions that the test was measuring their engineering aptitude negatively impacted their performance, suggesting that they are confirming this implicit threat that women have less ability than men in engineering. This study is parallel to a study by Stone et al. (1999) that examined black and white men's performance on a golf putting task. The men were randomly assigned to two conditions where they were told that the task was either a measure of athletic ability or of sports intelligence. The black men outperformed the white men when the task was characterized as measuring athletic ability, but the white men outperformed the black men when the task was characterized as measuring sports intelligence.

A study by Marchand and Taasoobshirazi (2012) on ST in high school physics compared men and women who were randomly assigned to one of three ST conditions including an explicit condition (men do better than women), implicit condition (no instructions regarding gender and performance), and nullified condition (no gender differences on the test). Results showed that the men outperformed the women on a set of physics problems in the implicit and explicit ST conditions, but that men and women performed similarly in the nullified condition. This suggested that simply being in a typical physics testing situation was enough to compromise women's performance, but that reminding students that both men and women are capable of doing well in physics removed any negative effects to women.

The present study is based on the study by Marchand and Taasoobshirazi (2012), but with chemistry as the subject of interest. Whereas physics has the largest gender gap in participation of any of the physical sciences (NSF 2015), we wanted to know if ST would play a role in chemistry where gender differences are not as prominent. Although we present studies above with more than two ST testing conditions, these are relatively rare. Occasionally, studies also include a reverse ST condition in which students are told that women perform better than men on the test (e.g., McIntyre et al. 2003). No research to date has included all four conditions, a gap the present study was designed to address.

Effects of ST on motivation

Recently, researchers have given increased attention to the mechanisms by which ST may exert influence in performance situations. Theoretical and empirical works recognize that ST effects are likely the result of a complex, multiple-influenced process (Doan and Hilpert 2009; Spencer et al. 2016). Mechanisms by which ST affects performance may include cognitive and working memory factors, physiological arousal, emotion, and motivational processes (Schmader et al. 2008; Shapiro 2011). For example, research has examined how goal orientation (Brodish and Devine 2009; Deemer et al. 2014), test-anxiety (Brodish and Devine 2009), and domain identification (Steinberg et al. 2012) mediate the effects of ST threat on performance and other outcomes. Further, ST effects may depend on degree of personal identification with the stereotype, task difficulty, and stereotype activation (Nguyen and Ryan 2008). Evidence for mediating and moderating factors has not always been consistent across studies and may suggest a context-dependency in terms of domain or developmental level (e.g., Flore and Wicherts 2015).

Research on gender and ST in science has only just touched on factors other than academic performance (Marchand 2015). In one study, an intervention aimed at having students express value beliefs ameliorated ST effects with females in undergraduate

physics courses (Miyake et al. 2010). Other works suggest that directly challenging stereotypes related to females in physics had a positive effect on student motivation and strategy use (Vollmeyer et al. 2009). In a study with undergraduate women in physics and chemistry, ST was associated with lower self-efficacy in chemistry and physics, but self-efficacy only mediated the relationship between ST and intent to pursue a science career in physics, but not chemistry (Deemer et al. 2014). With a few exceptions (e.g., Good et al. 2010) the emerging research on mechanisms associated with ST and performance has not necessarily examined effects of ST on motivation across multiple ST conditions.

Present study

There is insufficient research on the impact of ST on gender differences in performance in chemistry. Further, research detailing whether student motivation in chemistry varies in response to ST is scarce. This type of information could have important implications for understanding mechanisms of influence of ST in the domain of chemistry. Finally, there are no studies comparing four ST conditions; one study in physics compared three different ST conditions including an explicit condition, a nullified condition, and an implicit condition. The range of different factors under study and variance in the nature of effects within and across studies suggest that more domain-specific research is needed to identify the consistency with which ST effects are present within a domain and mechanisms of influence that may be domain-specific.

The goal of the present study was to extend the research on gender differences in chemistry and the research on ST in science by using a quasi-experimental design to compare the impact of ST on college women's chemistry performance across four ST conditions. This included an explicit ST condition (students were told men outperform women on the chemistry test), an implicit ST condition (students were not provided any information about the effect of gender on performance, but are in a traditional testing situation), a reverse ST condition (students were told women outperform men on the chemistry test), and a nullified condition (students were told that no gender differences in performance have been found on the test). We examined students' performance on a set of chemistry problems, their self-efficacy, and their test-anxiety across the four conditions. Self-efficacy and test-anxiety are considered two key motivational components that play a role in science achievement (Glynn et al. 2007).

Method

Participants

One hundred fifty three introductory level college chemistry students at a Midwestern university in the United States participated in the study. Seventy six students were male and 77 were female; approximately 64 % were Caucasian, 11 % were African–American, 7 % were Asian, and 1 % were Hispanic; the remaining participants either did not report their race, or marked 'other' or 'mixed' for their race. Students were randomly assigned to the four study conditions based on a cluster approach with recitation (lab) groups being the unit of assignment. Students spent 1 day each week in recitation, with approximately 15–30 students in each of six recitation groups. The random assignment of ST conditions to

recitation groups resulted in 42 students in the explicit ST group, 44 students in the reverse ST group, 36 students in the nullified ST group, and 31 students in the implicit ST group. An a priori power analysis using G Power (Faul et al. 2009) indicated that our sample size was larger than what G power recommended for a 2×4 MANOVA with three dependent variables, an effect size f^2 of .06, an alpha of .05, and a power of .80 (G Power recommended total sample size was n = 120).

Measures

Chemistry achievement

Students were given five chemistry problems to solve (Table 1). The first problem was derived from Heyworth (1999), the second, third, and fourth problems were from college chemistry textbooks (e.g., Reger et al. 1997; Zumdahl 1997), and the fifth problem was created by a chemistry professor based on stoichiometry problems in introductory level chemistry texts. The five problems were based on major topics in chemistry including stoichiometry, thermochemistry, and Gas Laws. The professor who taught the large lecture section of the course for the students confirmed that the students had not seen the problems previously, that the problems were at the appropriate level for the students, and that the students had learned the material assessed by the problems. Students were asked to solve the problems and show all of their work.

To obtain problem solution scores, students' responses on the five problems were scored by a chemistry teaching assistant at the university who was working on his doctorate in chemistry. A rubric was created by a chemistry professor and was used by the teaching assistant to score the problems. Each of the five problems was worth two points for a total of 10 points. Partial credit was provided consistently within and across problems for solving parts of the problems correctly.

Self-efficacy

The following seven items were derived from the Motivation Strategies for Learning Questionnaire (MSLQ) (Pintrich et al. 1993), revised to focus on the task at hand, and were given to students to assess their self-efficacy:

Table 1 Chemistry problems

1. A solution contains 1.10 g of sodium nitrate in 250.0 mL of solution. What is the molarity of this solution? (molar masses: Na = 22.99 g/mol, N = 14.01 g/mol, O = 16.00 g/mol)

2. Calculate the mass, in grams, of $Al(OH)_3$ (molar mass = 78.00 g/mol) formed by the reaction of 500. mL of .100 M NaOH with excess $Al(NO_3)_3$

3. When a 6.0-g sample of metal at 100.0 °C is added to 45.0-g of water at 22.60 °C, the final temperature of both the metal and the water is 32.81 °C. What is the specific heat of the metal? The specific heat of water is 4.184 J/g K

4. Mixtures of helium and oxygen are used in scuba diving tanks to help prevent "the bends." For a particular dive, 46 L He at 25 °C and 1.0 atm and 12 L O2 at 25 °C and 1.0 atm were pumped into a tank with a volume of 5.0 L. Calculate the partial pressure of each gas and the total pressure in the tank at 25 °C. (R = .0821 L atm/mol K)

5. What is the theoretical yield, in grams, of calcium oxide (molar mass = 56.08 g/mol) prepared from the reaction of 4.20 of calcium metal (molar mass = 40.08 g/mol) with oxygen gas (molar mass 32.00 g/mol)?

Even if the test is hard, I can do it.
I believe I can get an excellent grade on the test.
I believe I have the skills to do well on the test.
I expect to do well on the test.
I'm certain I can figure out how to do the most difficult problem on the test.
I can do the problems on this test if I don't give up.
I can do even the hardest problem on this test if I try.

Students responded to the items using a 7 point Likert scale that ranged from 1 = "not at all true of me" to 7 = "very true of me." The MSLQ has extensive evidence of reliability and construct validity. For our students, reliability as assessed by Cronbach's alpha was .95.

Test-anxiety

The following three items were derived from the Motivation Strategies for Learning Questionnaire (MSLQ) (Pintrich et al. 1993), revised to focus on the task at hand, and were given to students to assess their test-anxiety:

I am worried about failing this test.
I have an uneasy, upset feeling about taking this test.
I am nervous about how I will perform on this test.

Students responded to the items using a 7 point Likert scale that ranged from 1 = "not at all true of me" to 7 = "very true of me". The MSLQ has extensive evidence of reliability and construct validity. For our students, reliability as assessed by Cronbach's alpha was .94.

Procedures

Study materials were administered during the last week of class before the final exam. Students were given a packet with the study materials and were allowed to use a calculator to help them with their problem solving. Students were also given two equations including $PV = nRT$ and $q = mc\Delta t$. The instructions varied across the four conditions in just the following way:

Implicit ST condition

You will be given five chemistry problems to solve. These problems are based on chemistry material that you have already covered.

Explicit ST condition

You will be given five chemistry problems to solve. These problems are based on chemistry material that you have already covered. This test has shown gender differences with males outperforming females on the problems.

Nullified ST condition

You will be given five chemistry problems to solve. These problems are based on chemistry material that you have already covered. No gender differences in performance have been found on this test.

Reverse ST condition

You will be given five chemistry problems to solve. These problems are based on chemistry material that you have already covered. This test has shown gender differences with females outperforming males on the problems.

Before solving the chemistry problems and after the ST instructions, students were given the self-efficacy and test-anxiety survey items with the instructions "In order to better understand how you feel about this upcoming chemistry test, please respond to each of the following statements."

After students completed the chemistry problems, they were given four open-ended questions to answer. This included one question about the chemistry problems ("Was there anything that got in the way of, or interfered with, your performance?") and one question regarding students' academic or career path ("Do you plan to pursue a college major or a career in chemistry? Yes/No. Why or why not?). In addition, students were asked to answer two questions regarding their views of gendered capabilities ("Do you feel that men and women have the same mental capacity to achieve in chemistry? Please explain."), and gendered opportunities within the field of chemistry ("How about the same opportunities? Please explain."). Responses on the four open-ended questions were analyzed via a qualitative content analysis (QCA) approach. By reviewing these questions, the researchers hoped to better understand how the students perceived the relationship between gender, ST, and success in chemistry, as well as their overall sense of identification with the subject.

Students had approximately 55 min to complete the survey items and five chemistry problems. After the students completed the study and packets were collected, students were debriefed about the study.

Results

A 2×4 MANOVA was conducted to determine if differences in student performance, self-efficacy, and test-anxiety differed by gender (two groups: male and female) and ST condition (four groups: implicit ST, explicit ST, nullified ST, and reverse ST). Results of the MANOVA indicated a significant effect by gender Wilks' Lambda $= .935$, $F = 3.29$, $p = .03$. Neither the ST condition effect nor the interaction between ST condition and gender were significant. Thus results indicated that there were no differences by ST condition on the dependent variables and no gender differences across the ST conditions. Rather, there was a main effect for gender on the dependent variables. Follow up tests indicated those gender differences were on the self-efficacy ($p = .02$) and test-anxiety ($p = .01$) measures. The men had higher self-efficacy ($M = 32.71$, $SD = 10.23$) than the women ($M = 28.09$, $SD = 10.51$) and the men had lower test-anxiety ($M = 11.61$, $SD = 6.00$) than the women ($M = 14.38$, $SD = 5.96$). Significant differences between men and women were not found on performance on the chemistry problems. Tables 2 and 3

report descriptive statistics for men and women for the dependent variables across the ST conditions.

When self-efficacy and/or test-anxiety were used as covariates in an ANOVA, differences in performance on the chemistry problems were still not found. Therefore, even controlling for self-efficacy and assessment anxiety, differences between ST groups by gender were not found on the problems.

Responses on the four open-ended questions were analyzed and answers to the question: "Was there anything that got in the way of, or interfered with, your performance?", were largely practical in nature, with students citing their lack of preparation (ex: "unexpected test," "had not studied the information," "was not prepared") or the lack of test-taking aids they were provided (ex: "…Periodic table would have been helpful," "did not have the formula at hand"). Because none of the responses were directly related to students' views on gender or ST, this question was excluded from analysis. None of the responses indicated that the ST instructions impacted their performance. While the decision to focus analysis on only relevant materials was methodologically appropriate within our framework (Schreier 2012; MacQueen et al. 2009), the reason for exclusion of this question supports the quantitative findings indicating no impact of the ST conditions.

Qualitative Content Analysis (QCA) was utilized to assess the remaining three questions, as explained below. QCA developed as a method to retain the strengths of quantitative content analysis via "systematic text analysis" (Mayring 2000), while providing the flexibility to explore qualitative data with a more contextualized analytical lens (Krippendorff 2004; Schreier 2012). QCA is appropriate when a strong body of literature about a topic already exists and the research goal is to describe trends or themes (Cho and Lee 2014). This is opposed to a study seeking to generate a new theory from the data, a necessary component that differentiates Grounded Theory from the descriptive goals of this analysis (Glaser and Strauss 1967; Goulding 2002). Because our goal was not to develop alternative theories around stereotype threat, but to better understand our results within the existing theory, QCA was determined to be the most appropriate method.

As outlined by Elo and Kyngäs (2008), three phases hold across the majority of QCA definitions: preparing, organizing, and reporting. In the preparation phase, the researcher needs to select the unit of analysis (Elo and Kyngäs 2008; Schreier 2012). For this study, because individual question responses were brief, data were segmented by question (i.e. each response was analyzed and categorized independently). We adopted an inductive model (Graneheim and Lundman 2004, Elo and Kyngäs 2008; Cho and Lee 2014; Schreier 2012), or a model wherein the categories within the data are allowed to emerge via open coding, rather than imposed in advance.

Table 2 Descriptive statistics for women across groups for the three dependent variables

	Implicit ST condition			Explicit ST condition			Nullified ST condition			Reverse ST condition		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
Performance	3.76	2.39	17	3.78	2.27	20	2.91	2.37	17	4.50	3.05	23
Self-efficacy	29.47	10.90	17	27.50	10.98	20	29.94	10.47	17	26.22	10.18	23
Test anxiety	14.59	5.23	17	14.75	6.79	20	14.76	6.17	17	13.61	5.88	23

Table 3 Descriptive Statistics for Men across Groups for the Three Dependent Variables

	Implicit ST condition			Explicit ST condition			Nullified ST condition			Reverse ST condition		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
Performance	3.39	2.57	14	4.75	3.26	22	3.21	3.06	19	4.07	2.80	21
Self-efficacy	28.71	10.66	14	34.05	10.23	22	30.47	9.82	19	36.00	9.61	21
Test anxiety	13.93	5.94	14	11.32	6.71	22	12.32	5.23	19	9.76	5.91	21

After the unit of analysis was identified, one researcher identified trends and formulated questions about the data. Most, but not all, responses were structured with a directional statement (Yes/No/other), followed by an elaboration that sought to defend, explain, or verify the directional statement. For instance, a student might say "Yes. I will have a career in chemistry because it's my favorite subject." The overall endorsement of a "yes/no/ other" position is reinforced or contradicted via the second, explanatory statement. This second portion of the responses provided rich insight into the ways students interpreted their experiences studying chemistry. Because of this, it was determined that responses would be categorized across two categories.

The first category considered the direction of a student's response. Although the questions were phrased as Yes/No, with the opportunity to elaborate, students often began a response by explicitly saying "Yes," or "No," but their elaborations either mitigated or reversed their explicit response (ex: "Yes [women have the same opportunities], but I see how sometimes opportunities are limited for women"). Sometimes responses explicitly stated that the student was unsure, or that their answer was contingent on a variety of factors. To best represent this diversity, responses were categorized as "Yes," "No," "Unsure," or "Mixed," and this determination was made with consideration to what the student explicitly indicated as well as the content of their explanation.

The second aspect of each response, or category, that we reviewed was the reasoning, or explanations, provided for the "yes/no/unsure/mixed" responses. During open coding, several classifications were noted within student explanations. These included references to innate skills or intellect (ex: "Sure, [mental capacity] depends on your brain function type"), the importance of effort (ex: "Yes, I think that mental capacity is not [b]ased of off gender, but rather how much an individual is willing to work"), and social training and stereotypes, (ex: "…it has always been stereotyped that men are better in science than women), as well as statements about group and individual preference, among numerous others.

After the initial categories and underlying classifications were developed and abstracted, one author constructed definitions of each for coding and categorized each response based on these definitions. To ensure reliability, a codebook was developed to allow a second researcher to categorize responses independently. Definitions for categories and classifications included guidelines for when to apply and not apply a code as well as examples. Agreement was initially high across the first category (yes, no, unsure, mixed: 95 % agreement), but low across the second category (explanations). Based on Krippendorff's (2004) guidelines, both categories were assessed independently so as to avoid artificial inflation of agreement via an index score. The codebook was further refined over two rounds of independent scoring, discussion, and revision, with a final round of

independent coding demonstrating strong reliability for category one ($\alpha = .926$; 95.9 % agreement; 4 classifications) and acceptable reliability for category two ($\alpha = .731$; 75.4 % agreement; 19 classifications) for the "tentative conclusions" sought in this study (Krippendorff 2004). All data were then recoded by the primary coder using the revised definitions (MacQueen et al. 1998). Responses across category one, by question, by gender, are represented in Table 4. The top three explanatory classifications offered by question, by gender, are included in Table 5.

In the context of the quantitative data produced via this study, the most interesting trend identified within the qualitative responses is the overwhelming lack of domain identification among participating students. Domain identification has been considered an important element within the ST literature, as students who identify with a given domain are thought to be more affected by ST than students who do not identify as such. This is both a theoretical proposition (Steele 1997) as well as an empirically supported position (Aronson et al. 1999; Keller 2007). In our study, more than 85 % of respondents indicated that they would not pursue a career in chemistry. For women, the rate was slightly higher (88 %) and for men, slightly lower (82 %).

When asked to elaborate, nearly 50 % of respondents focused on explicit dislike of chemistry or preference for another subject, while another 23 % explained chemistry as simply a hurdle to their goal major or degree. When elaborating, students often described chemistry as "tedious," stating that "it doesn't interest me," or that they "…only need it to get [their] degree." The fact that most of these students likely enrolled due to degree requirements, but were at pains to distance themselves from the field, may supplement our understanding of domain identification within the group, including the fact that these students may have been less likely to be impacted by ST conditions.

In addition, responses to the other questions indicated that the women were more likely than the men to report gender differences in mental capacity and opportunity. In direct response to the question "Do you feel that men and women have the same mental capacity to achieve in chemistry? Please explain," women did answer "no" more often than their male counterparts (10 vs. 4 %). These women also explained their responses, regardless of

Table 4 Yes, no, unsure, and mixed themed responses by question by gender

Gender	Mixed (%)	Yes (%)	No (%)	Unsure (%)
Question 1: Do you plan to pursue a college major or a career in chemistry? Yes/No. Why or why not?				
Men	1.47	14.71	82.35	1.47
Women	1.69	8.47	88.14	1.69
All	1.54	11.54	85.38	1.54
Question 2: Do you feel that men and women have the same mental capacity to achieve in chemistry? Please explain				
Men	2.94	92.65	4.41	.00
Women	.00	87.93	10.34	1.72
All	1.54	90.77	6.92	.77
Question 3: How about the same opportunities? Please explain				
Men	4.69	56.25	32.81	6.25
Women	9.26	40.74	44.44	5.56
All	7.32	50.41	36.59	5.69

Table 5 Top Three Response Themes by Gender by Question

Question	1st	2nd	3rd
Male top frequencies (%). Examples are from either gender			
Q1	Dislike of Chemistry (43.86)	Chemistry as Hurdle to Goals (19.30)	Chemistry Aptitude as a Fixed, Individual Characteristic (15.79)
Sample responses	No. I hate chemistry.//No. I don't like chemistry. I like bio better	This is just a requirement to get into my college.//No. Chemistry is required for my major but I am not interested in it as a career	No. Chemistry is not my strong suit.//No. God no. I am not chemistry inclined
Q2	Citing a Lack of Perceived Gender Difference (42.11)	Chemistry Aptitude as Effort-Based (29.82)	Chemistry Aptitude as a Fixed, Individual Characteristic (12.28)
Sample responses	Yes, everyone is equal.//Yes, because we are all human.//Yes, gender does not matter	Yes I believe anyone can achieve anything with proper ambition.//Of course, I don't believe it's the gender, it's the effort to learn is what separates people, men and/or women	Yes. Some people have math/chemistry brains and some don't. I don't think it has anything to do with gender.
Q3	Citing a Lack of Perceived Social Gender Differences (27.08)	Citing Social Stereotypes (22.92)	Deferring to Social Observations[a] (12.50)
Sample responses	Yes, we are all offered the same resources.//Yes, men and women are given the same tests/classes	No. women have to do more to prove themselves. Many men have a problem taking women seriously.//I feel like women are pushed more toward jobs other than chemistry related ones and do not get the same opportunities	Men have more opportunities because more men are chemists or chemical engineers, but are closely followed by women now!
Female top frequencies (%). Examples are from either gender			
Q1	Dislike Chemistry (29.17)	Chemistry as Hurdle to Goals (27.08)	Chemistry Aptitude as a Fixed, Individual Characteristic (22.92)
Sample responses	See above	See above	See above
Q2	Chemistry Aptitude as Effort-Based (28.85)	Citing a Lack of Perceived Gender Difference (21.15)	Chemistry Aptitude as a fixed, Gender-Based Characteristic (15.38)

Table 5 continued

Question	1st	2nd	3rd
Sample responses	See above	See above	No, because both genders think differently.//Yes, even though memorization usually seems to come easier to men
Q3	Citing Social Stereotypes (34.78)	Citing a Lack of Perceived Social Gender Difference (21.74)	Chemistry Access as Context Dependent (19.57)
Sample responses	See above	See above	I think [opportunity] depends on the field because some fields people believe one gender can do it better than the other

[a] Responses that took a position on a given question, but avoided explanation or theorizing, were placed in this category. In these responses, the student would point to "everyday observations," as support, often without considering the context from which their observations derived

answering Yes or No, by referencing innate differences between genders (15 % women vs. 0 % men). When asked whether women and men had the same opportunities, however, men were more likely to indicate that opportunities were equal (56 vs. 41 %), and less likely to offer explanations that acknowledged social bias or stereotypes (25 % of men vs. 40 % of women). This suggests that a minority of women may be more likely to believe in innate differences between men and women, but a larger proportion demonstrated more awareness of gender differences in opportunities in science.

The most popular explanations among women in response to the mental capacity question were those focusing on the importance of effort over innate skill (29 %) with an equal proportion of men endorsing this idea (30 %). Women's second most frequent response theme (21 %), and men's most frequent theme (42 %), were responses that focused on a lack of identifiable mental difference between men and women (ex: "everyone is equal," "there isn't anything preventing a man or woman in terms of learning," or "we are all human"). These responses tended to stop short of theorizing where gender-based performance differences might have arisen. Overall, 50 % of female responses to the question about mental capacity fell into one of these two categories (focusing on effort or a lack of discernable differences). This number, as well as overwhelming endorsement of equal mental capacity across genders (88 %), suggests that the women (and men) in this study may be more likely to view their intellectual competence as "malleable" as opposed to "fixed." Studies have found that viewing intelligence as increasing via effort can serve as a protective factor against the theorized "burden" of "confirming cultural stereotypes impugning their intellectual and academic abilities," a key component to ST theory (Aronson et al. 2002; Good et al. 2003). These results, along with the findings that women and men in the study (a) distanced themselves from chemistry and (b) endorsed the idea that men and women had equal mental capacities, provide interesting insights into the relative weight that the female and male students in this study gave to social stigma, effort, and biology. Specifically, these qualitative findings provide additional insight as to why this group of students did not exhibit the predicted ST results across the four conditions.

Discussion

This is the first study to compare the impact of four ST conditions on the performance, self-efficacy, and test-anxiety of chemistry students. We found no differences by ST condition or difference for ST condition by gender; rather we found a significant main effect for gender with men reporting higher self-efficacy and lower test-anxiety. An analysis of responses on open-ended questions asking students about their intent to major in chemistry, beliefs regarding barriers to their achievement on the test, and gender differences in opportunities and mental capacity to achieve in chemistry provided insight into the quantitative findings.

Our results indicated that the ST instructions did not impact the students' performance on the chemistry problems. We know that the gender gap in chemistry is closing for undergraduate chemistry majors (Matson 2013). It is possible that the negative effects of ST found in physics, engineering and mathematics, which are more male-dominated, are not an issue in chemistry. Studying ST in biology, where the gender discrepancy in enrollment favors women, may provide additional insight into the domains in which ST may have less of an effect. It is also important for researchers to determine when the effects of ST on performance may begin to emerge and peak. It may be that although we did not

find an effect at the college level, that differences may be found at the high school level. On the other hand, it may be that ST plays a role in more advanced chemistry courses with chemistry majors where we see more of a disparity in gender enrollment and where there is greater domain identification. We know that individuals who highly identify with a domain are more likely to be impacted by stereotype threat (Aronson et al. 1999). As such, research should compare the impact of ST across various levels of science courses where students exhibit different levels of motivation and domain identification.

There were gender differences with men reporting higher self-efficacy and lower test anxiety across all of the ST conditions. This is consistent with a large body of research in science education illustrating that women have lower self-efficacy and higher test-anxiety than men in their introductory-level college science courses (Cavallo et al. 2004; Glynn et al. 2009). However, these differences did not translate to differences in performance,[1] calling into question, at least for this group of students, the stereotype threat explanation in which negative motivational states interfere with cognitive performance during testing. Although these gender differences in self-efficacy and test-anxiety did not result in gender differences in performance for this group of students, it may impact the women's performance later down the road in more advanced chemistry courses. In addition, these differences may be enough to keep women from persisting and participating in chemistry. Thus these differences in self-efficacy and test-anxiety should be addressed.

Although not a large number of women, the qualitative findings indicated that more women than men believed in innate differences between men and women. In addition, women were more likely than men to be expectant of gender differences in opportunities in science. These could have contributed to women's lower self-efficacy and higher test-anxiety.

Although the chemistry problems that were given to the students were confirmed to be at the appropriate level for the students, the mean score on the chemistry problems for all the students was 5.2 out of 10 points, indicating that students scored at about 50 %. Therefore, the problems were not too easy or too difficult. However, we know little from the ST research about how ST impacts student performance on tasks beyond basic classroom testing. What is not yet known is how males and females perform in non-testing scenarios, such as authentic class projects, which may reflect ST effects.

A growing body of research has begun to explore how motivational and cognitive variables exasperate or ameliorate the impact of ST, though this research is just emerging in the sciences. Some of the important variables that have been implicated as important contributors to ST have included self-regulation, goal orientation, ST endorsement, and domain-identification, to name a few (Schmader et al. 2008; Smith 2004; Steinberg et al. 2012). Our quantitative findings indicate that men and women differ in their self-efficacy and test-anxiety, two important motivational variables. The qualitative analysis in the present study suggested that men and women did differ with respect to their perceptions of male and female opportunities and abilities in science, as well as their degree of identification with chemistry. These exploratory analyses, while providing some insight and qualification of the quantitative results, suggest directions that could be explored in future research. For example, a structural equation model with multiple cognitive and motivational variables can provide information about how motivational variables may mediate the ST-performance relationship. Cluster analysis can be used to obtain motivational and cognitive profiles to determine which are most conducive in protecting one from the

---

[1] The correlation between self-efficacy and performance for all 153 students was $r = .43$, $p < .001$; the correlation between test-anxiety and performance was $r = -.40$, $p < .001$.

negative effects of ST (Marchand 2015). Clearly, additional research is needed to determine the prevalence of ST effects in chemistry and whether these effects are less robust than in other STEM domains and to explore the complex mechanisms involving context, domain, and individual differences that might lead to ST effects in performance (Doan and Hilpert 2009; Spencer et al. 2016).

Conclusion

Prior studies in mathematics and physics have found that ST affects women's academic performance. This was the first study to examine ST using four different conditions in chemistry at the post-secondary level. Even though we found no differences in gender across the four ST conditions, we did find that, across all four conditions, women had lower self-efficacy and higher anxiety in chemistry. These differences, however, did not translate to differences in performance in this introductory level chemistry course on chemistry problems that are typical of tests and assignments in the course, perhaps due to low domain identification among the participating students. These gender differences, however, could eventually impact participation and later achievement. We would like to see more studies examining ST across different domains of science and across different levels of study. We recommend that future research on ST examine more advanced chemistry students and also biology students. We also recommend studying how motivational, cognitive, and contextual variables may exasperate or ameliorate the impact of ST in science.

References

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*(2), 113–125.

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.

Bell, A. E., Spencer, S. J., Iserman, E., & Logel, C. E. (2003). Stereotype threat and women's performance in engineering. *Journal of Engineering Education, 92*(4), 307–312.

Blickenstaff, C. J. (2005). Women and science careers: leaky pipeline or gender filter? *Gender and Education, 17*(4), 369–386. doi:10.1080/09540250500145072.

Brodish, A. B., & Devine, P. G. (2009). The role of performance–avoidance goals and worry in mediating the relationship between stereotype threat and performance. *Journal of Experimental Social Psychology, 45*(1), 180–185. doi:10.1016/j.jesp.2008.08.005.

Cavallo, A. M., Potter, W. H., & Rozman, M. (2004). Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors. *School Science and Mathematics, 104*(6), 288–300.

Chen, X. & Soldner, M. (2013). STEM Attrition: College Students' Paths into and out of STEM Fields. Statistical Analysis Report. NCES 2014-001. National Center for Education Statistics.

Cho, J. Y., & Lee, E. H. (2014). Reducing confusion about grounded theory and qualitative content analysis: Similarities and differences. *The Qualitative Report, 19*(32), 1–20.

Deemer, E. D., Smith, J. L., Carroll, A. N., & Carpenter, J. P. (2014a). Academic procrastination in STEM: Interactive effects of stereotype threat and achievement goals. *The Career Development Quarterly, 62*(2), 143–155. doi:10.1002/j.2161-0045.2014.00076.x.

Deemer, E. D., Thoman, D. B., Chase, J. P., & Smith, J. L. (2014b). Feeling the threat stereotype threat as a contextual barrier to women's science career choice intentions. *Journal of Career Development, 41*(2), 141–158.

Desouza, J. M. S., & Czemiak, C. M. (2002). Social implications and gender differences among preschoolers: Implications for science activities. *Journal of Research in Childhood Education, 16*, 175–188. doi:10.1080/02568540209594983.

Doan, L. & Hilpert, J. (2009). *A Qualitative Investigation of Stereotype Threat Activation and Mediation.* Presented at American Educational Research Association Annual Meeting, San Diego, CA

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107–115.

Enman, M., & Lupart, J. (2000). Talented female students' resistance to science: An exploratory study of post-secondary achievement motivation, persistence, and epistemological characteristics. *High Ability Studies, 11*(2), 161–178. doi:10.1080/13598130020001205.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology, 53*(1), 25–44.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory.* London: Weidenfeld and Nicholson.

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2007). Nonscience majors learning science: A theoretical model of motivation. *Journal of Research in Science Teaching, 44*, 1088–1107.

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2009). Science motivation questionnaire: Construct validation with nonscience majors. *Journal of Research in Science Teaching, 46*(2), 127–146.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*(6), 645–662.

Good, J. J., Woodzicka, J. A., & Wingfield, L. C. (2010). The effects of gender stereotypic and counter-stereotypic textbook images on science performance. *The Journal of Social Psychology, 150*(2), 132–147. doi:10.1080/00224540903366552.

Goulding, C. (2002). *Grounded theory: A practical guide for management, business and market researchers.* Newbury Park: Sage.

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today, 24*(2), 105–112.

Greene, B. A., & DeBacker, T. K. (2004). Goal and orientations toward the future: Links to motivation. *Educational Psychology Review, 16*, 91–120. doi:10.1023/B:EDPR.0000026608.50611.b4.

Heyworth, R. M. (1999). Procedural and conceptual knowledge of expert and novice students for the solving of a basic problem in chemistry. *International Journal of Science Education, 21*(2), 195–211. doi:10.1080/095006999290787.

Hill, C., Corbett, C., & St Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics.* Washington, DC: American Association of University Women.

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle teaching stereotype threat as a means of improving women's math performance. *Psychological Science, 16*(3), 175–179.

Johnson, H. J., Barnard-Brak, L., Saxon, T. F., & Johnson, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education, 80*(2), 137–149. doi:10.1080/00220973.2011.567312.

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology, 77*(2), 323–338.

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology, 43*(5), 825–832. doi:10.1016/j.jesp.2006.08.004.

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research, 30*(3), 411–433.

MacQueen, K. M., McLellan, E., Kay, K., & Milstein, B. (1998). Codebook development for team-based qualitative analysis. *Cultural Anthropology Methods, 10*(2), 31–36.

Marchand, G.C. (2015). *Stereotyp threat and female performance in secondary physics courses.* Presented at the American Association of Physics Teachers, San Diego, CA

Marchand, G. C., & Taasoobshirazi, G. (2013). Stereotype threat and women's performance in physics. *International Journal of Science Education, 35*(18), 3050–3061. doi:10.1080/09500693.2012.683461.

Matson, J. (2013, April). Women are earning greater share of STEM degrees, but doctorates remain gender skewed. *Scientific American, 308*(5). Retrieved from http://www.scientificamerican.com/article/women-earning-greater-share-stem-degrees-doctorates-remain-gender-skewed/

Mattern, N., & Schau, C. (2002). Gender differences in science attitude-achievement relationships over time among white middle-school students. *Journal of Research in Science Teaching, 39*(4), 324–340. doi:10.1002/tea.10024.

Mayring, P. (2000). *Qualitative content analysis. forum qualitative sozialforschung/forum: Qualitative social research, 1*(2). http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204.

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology, 39*(1), 83–90. doi:10.1016/S0022-1031(02)00513-9.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*(6008), 1234–1237.

National Science Foundation. (2015). *Women, minorities, and persons with disabilities in science and engineering.* Retrieved from NSF website: http://www.nsf.gov/statistics/2015/nsf15311/digest/theme4.cfm#minoritybachelors

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.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29*(6), 782–789. doi:10.1177/0146167203252810.

Olson, S., & Riordan, D. G. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Executive Office of the President: Report to the President.

Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53*(3), 801–813.

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues, 57*(1), 55–71.

Reger, D. L., Goode, R. S., & Mercer, E. E. (1997). *Chemistry: Principles and practice.* Belmont, CA: Brooks Cole.

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., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 336.

Schreier, M. (2012). *Qualitative content analysis in practice.* London: Sage Publications.

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.

She, H. C. (2001). Different gender students' participation in the high-and low-achieving middle school questioning-orientated biology classrooms in Taiwan. *Research in Science & Technological Education, 19*(2), 147–158. doi:10.1080/02635140120087696.

Shin, N., & McGee, S. (2002, November). *The influence of inquiry-based multimedia learning environment on specific problem-solving skills among ninth grade students across gender differences.* Paper presented at Association for Educational Communications and Technology, Dallas, TX.

Smith, J. L. (2004). Understanding the process of stereotype threat: A review of mediational variables and new performance goal directions. *Educational Psychology Review, 16*(3), 177–206. doi:10.1023/B:EDPR.0000034020.20317.89.

Smith, J. L., & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles, 47*(3–4), 179–191. doi:10.1023/A:1021051223441.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology, 67*, 415–437.

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

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 and Applied Social Psychology, 34*(6), 534–543. doi:10.1080/01973533.2012.727319.

Stieff, M., Ryu, M., Dixon, B., & Hegarty, M. (2012). The role of spatial ability and strategy preference for spatial problem solving in organic chemistry. *Journal of Chemical Education, 89*(7), 854–859.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology, 77*(6), 1213–1227. doi:10.1037/0022-3514.77.6.1213.

Tenenbaum, H. R., & Leaper, C. (2003). Parent–child conversations about science: The socialization of gender inequities? *Developmental Psychology, 39*(1), 34–47. doi:10.1037/0012-1649.39.1.34.

Vollmeyer, R., Püttmann, A., & Imhof, M. (2009). How to improve women's performance in physics through stereotype threat. In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 1471–1476).

Wang, X. (2013). Modeling entrance into STEM fields of study among students beginning at community colleges and four-year institutions. *Research in Higher Education, 54*(6), 664–692. doi:10.1007/s11162-013-9291-x.

Zumdahl, S. S. (1997). *Chemistry*. Boston, MA: Houghton Mifflin.