

Stereotype threat and gender differences in statistics

Gita Taasoobshirazi ^{1*}, Ordene Edwards ², Bowen Eldridge ¹

¹School of Data Science and Analytics, Kennesaw State University, Kennesaw, GA, USA

²Department of Psychological Sciences, Kennesaw State University, Kennesaw, GA, USA

*Corresponding Author: gtaasoo@kennesaw.edu

Citation: Taasoobshirazi, G., Edwards, O., & Eldridge, B. (2023). Stereotype threat and gender differences in statistics. *Contemporary Mathematics and Science Education*, 4(1), ep23014. <https://doi.org/10.30935/conmaths/13064>

ABSTRACT

Stereotype threat (ST) has been extensively explored as an explanation for gender disparities in achievement and participation in mathematics. However, there is a lack of research evaluating ST in statistics. The present study evaluated the impact of ST on gender differences in student performance, self-efficacy, and anxiety in statistics using a four-group, quasi-experimental design. Specifically, 102 elementary statistics students at a university in the Southeast United States 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 ST condition. Results indicated that there were no gender differences by ST condition in statistics self-efficacy, test anxiety, and performance. Analyses of student responses to open-ended questions indicated that females were more likely than males to report that they had fewer opportunities to achieve in statistics. Implications of our findings and suggestions for future research are discussed.

Keywords: stereotype threat, statistics, gender, achievement, motivation

Received: 24 Jan. 2023 ♦ Accepted: 10 Mar. 2023

INTRODUCTION

There is a large amount of research that attempts to explain gender disparities in K-16 mathematics achievement and participation (Pilotti, 2021). Explanations for why outcomes favor males have included motivation, strategy use, metacognition, social support, socio-economic status, and stereotype threat (ST) (Brown et al., 2016; Carr & Jessup, 1997; Jacobs, 2005; Pilotti, 2021; Witherspoon & Schunn, 2020). Much less research has evaluated gender disparities in performance and participation in statistics, biostatistics, and data science despite reports of gender disparities in participation in statistics related occupations. For example, a recent report documented that less than 17% of data analytics jobs are filled by women (Boas, 2020). This lack of research in statistics is, in part, because students typically do not get the opportunity to take a statistics course until late in high school and/or in college (Yates et al., 2021). A second explanation is that in many progress reports on degrees awarded, mathematics and statistics outcomes are combined (NSF, 2017). However, statistics is different from mathematics in its approach and application and should be evaluated as its own area of study (Rossman et al., 2006). Specifically, mathematics follows a rigid theorem and proof structure whereas statistics is a discipline in which real-life data is handled. Further, in mathematics, space, measures, and structures in their rudimentary form are considered, whereas in statistics raw data is collected, sorted, interpreted, and represented. Finally, mathematics is a subject of more

absolute conclusions whereas statistics predictions are uncertain and vary with context (Mathnasium, 2022).

Understanding the role of gender in statistics performance and participation is essential given that, in the last decade, there has been a great need for talented individuals who can organize and analyze the enormous amount of data that is being collected (Legaki et al., 2020).

This continuous inflow of data has put statistics, biostatistics, and data science at the forefront of desired majors and careers and has increased the need for more analysts (Stern et al., 2020). For example, Forbes (2020) listed statistics as one of the 15 most valuable college majors. In 2020, Glassdoor (2020) listed data scientist as one of the best jobs in America. In addition, there is a need for diversity in analysts. In her book, *Invisible women: Exposing data bias in a world designed for men*, Criado-Perez (2019) describes the disadvantages to women that result from biased data analyses that is from a primarily male perspective. Criado-Perez (2019) explains that because data is mainly analyzed by men and for men, this creates systematic bias and discrimination against women, causing inequities in industry, medical, and technological data analyses and applications. When data is unbalanced and excludes certain groups, models are biased. To support bringing more women into analytics related occupations, the Women in Data Science (WiDS) organization has put forth the initiative of "30 by 30" or having 30 percent of data scientists be women by the year 2030 (from the 2020 documented 17% of WiDS). While there is a clear gap in women's participation in statistics and data science related occupations, little is known about why these disparities in participation exist. It is unclear

whether the same causes for gender disparities in mathematics, such as differences in motivation, hold true in statistics. We cannot assume this to be the case. Research is needed to study the contributing factors to gender disparities in statistics, specifically.

STEREOTYPE THREAT

ST research developed from studies of racial differences in performance on standardized tests (Steele & Aronson, 1995). ST is the notion that beliefs held about a particular group may cause the confirmation of the judgment about one's group, and in turn, impact learning and performance (Johnson et al., 2012).

Much of what we know about gender and ST comes from the research in mathematics (Spencer et al., 1999). Different combinations of experimental or quasi-experimental conditions have been implemented to evaluate the impact of ST on gender disparities in mathematics. ST has been studied in conditions where the threat is made implicit (e.g., just being in a typical mathematics testing situation), explicit (e.g., students are told men perform better than women on a test), or nullified (e.g., equating the groups by telling students there are no gender differences) (Smith & White, 2001, 2002). Usually, implicit or explicit ST conditions are compared with a nullified condition (e.g., O'Brien & Crandall, 2003). For example, Spencer et al. (1999) compared college-level men and women's mathematics performance across two ST 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. The authors concluded that women underperformed in the implicit condition because of the existing implicit stereotype that women are less competent than men in mathematics. This is a consistent finding and interpretation across the research on ST in mathematics (O'Brien & Crandall, 2003; Quinn & Spencer, 2001).

We found one study that evaluated the impact of ST on women's performance and anxiety in statistics. Kapitanoff and Pandey (2017) examined whether gender of instructor was related to college, elementary-statistics students' anxiety, performance, and ST endorsement. For men, anxiety and performance were not linked to gender of their instructor; also, the men's anxiety was not linked to their ST endorsement. For women, having a female instructor initially resulted in worsened performance that disappeared over the term. The women's anxiety was linked to their ST endorsement. A limitation of the study was that the authors measured ST endorsement in mathematics rather than statistics. In addition, the anxiety questions measured anxiety in general rather than statistics specific anxiety.

The research on ST conditions on gender differences in mathematics has tended to focus almost exclusively on performance as an outcome. ST researchers have urged the study of how ST may impact motivation (Fogliati & Bussey, 2013). Research by Davies et al. (2002) found that women exposed to ST reported less interest in pursuing college majors and careers in quantitative domains (Davies et al., 2002). Fogliati and Bussey (2013) found that college women in a stereotyped condition (students were told men outperform women on a test) were less motivated than women in a nullified condition (students were told men and women perform equally well on the test) to use feedback to

revise and improve their mathematics work. There is a dearth of research evaluating the impact of ST on specific motivational constructs such as self-efficacy and anxiety, both of which have shown to have a significant impact on mathematics achievement and participation (Hiller et al., 2021). There is no research on this in statistics. Below, we review the extant research on gender differences in statistics achievement and motivation.

GENDER DIFFERENCES IN STATISTICS ACHIEVEMENT AND MOTIVATION

A search for peer-reviewed journal articles on gender differences in statistics achievement led to conflicting findings. At the college level, some research documented no gender differences in course grade among students taking undergraduate statistics (Buck, 1985; Es & Weaver, 2018; Woehlke & Leitner, 1980), others found results in favor of women (Charles, 1987), while others found results in favor of men (Susbiyanto et al., 2019). A meta-analysis of 18 empirical studies showed that men outperformed women on college statistics course examinations, but women outperformed men in the course overall (Schram, 1996). Similar varying findings were found for high school statistics students. For instance, Saidi and Siew (2019) found that men outperformed women on a test of central tendency; Batanero et al. (2003), in contrast, found no gender differences in high school students' understanding of central tendency.

The research on gender differences in motivation, and specifically self-efficacy and anxiety, in statistics is limited, dated, and has shown varying findings. Onwuegbuzie (1995) and Stroup and Jordan (1982) found that college women experienced higher levels of statistics anxiety than men. However, other studies did not find significant differences between college men and women in their statistics anxiety (Baloglu, 2001; Cruise et al., 1980). There is a lack of research evaluating gender differences in self-efficacy in statistics. This is an area that needs exploration given the significant impact of self-efficacy on achievement and participation in mathematics (Evans et al., 2021).

PRESENT STUDY

The present study is the first to examine the impact of ST on performance in statistics using four experimental conditions including an explicit ST condition (students are told men outperform women on the statistics test), an implicit ST condition (students are not provided any information about the effect of gender on performance, but are in a traditional testing situation), a reverse ST condition (students are told women outperform men on the statistics test), and a nullified condition (students are told that no gender differences in performance have been found on the statistics test). This study is unique in that it considers a reverse ST condition. In addition, it is the first to study gender differences in self-efficacy and anxiety in statistics under ST conditions.

METHOD

Participants

102 introductory level college statistics students at a Southeastern University in the United States participated in the study. 39 students identified as male and 63 identified as female; approximately 54% were

Table 1. Statistics problems

No	Question				
	Given this five-number summary, are there any lower and/or upper outliers?				
1	Minimum	Q1	Median	Q3	Maximum
	2	36	44	51	107
2	Population parameters of some bell-shaped distribution include a mean of 44 & an SD of 3.7. What are cut-off values that define middle 95% of data values?				
3	Given the same distribution in question 2, what is the z-score for the data value 48?				
4	The standard deviation of a given sample is 19.3. What is the sample variance?				

Caucasian, 25% were African American, 9% were Asian, 8% were Hispanic, and 4% identified as Other. The gender and race distribution for the statistics classes is equivalent to the undergraduate population at the university. Four classes, taught by two different instructors, each with approximately 35 students enrolled, participated. Students were randomly assigned to the four study conditions based on a cluster approach using the classroom as the unit of assignment. The random assignment of ST conditions to classroom resulted in 27 students in the explicit ST group, 33 students in the reverse ST group, 20 students in the nullified ST group, and 22 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 2x4 multivariate analysis of variance (MANOVA) with three dependent variables, an effect size f^2 of .15, an alpha of .05, and a power of .80 (G*Power recommended total sample size was $n=56$).

Measures

Statistics achievement

Students were asked to solve four statistics questions (Table 1). The questions, which were textbook problems (Navidi & Monk, 2019) with the values altered, tested descriptive statistics, z-scores, and the normal curve. The two professors teaching the courses confirmed that the students had not seen the problems previously, that the problems were at the appropriate level, and that the students had learned the material assessed by the problems.

Self-efficacy

The following seven items were derived from the motivation strategies for learning questionnaire (MSLQ) (Pintrich et al., 1993), revised to focus on statistics, and were given to students to assess their study-condition specific self-efficacy:

- 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." MSLQ has extensive evidence of reliability and construct validity. For our students, reliability as assessed by Cronbach's alpha was .94.

Test anxiety

The following three items were derived from MSLQ (Pintrich et al., 1993), revised to focus on statistics, and were given to students to assess their study-condition specific 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." MSLQ has extensive evidence of reliability and construct validity. For our students, reliability as assessed by Cronbach's alpha was .92.

Procedures

Study materials were administered during the seventh week of classes to ensure that students were exposed to the material on the test and to give them full exposure to the statistics learning environment. Students were given a packet with the surveys and statistics questions. The instructions varied across the four conditions in just the following way:

1. **Implicit ST condition:** You will be given four statistics problems to solve. These problems are based on statistics material that you may have already covered.
2. **Explicit ST condition:** You will be given four statistics problems to solve. These problems are based on statistics material that you may have already covered. This test has shown gender differences with males outperforming females on the problems.
3. **Nullified ST condition:** You will be given four statistics problems to solve. These problems are based on statistics material that you may have already covered. No gender differences in performance have been found on this test.
4. **Reverse ST condition:** You will be given four statistics problems to solve. These problems are based on statistics material that you may have already covered. This test has shown gender differences with females outperforming males on the problems.

Before solving the statistics problems and after ST instructions, students were given the self-efficacy and test-anxiety items with the instructions "in order to better understand how you feel about this upcoming statistics test, please respond to each of the following statements." Students had approximately one hour to complete the survey questions and statistics problems. After students completed all the surveys and statistics questions, we asked them the following questions:

1. Do you plan pursue a college major or a career in statistics? (O Yes O No).
2. Do you feel that men and women have the same mental capacity to achieve in statistics? Please explain. How about the same opportunities? Please explain.

Our research was conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with local

Table 2. Descriptive statistics by stereotype threat condition and gender

Stereotype threat condition		Males: Mean (SD)	Females: Mean (SD)
Reverse: 11 males & 22 females	Performance	3.05 (1.31)	3.14 (1.14)
	Self-efficacy	36.45(7.95)	34.58(8.97)
	Anxiety	10.36(4.61)	9.86(5.89)
Explicit: 3 males & 19 females	Performance	3.31(1.31)	3.35 (1.28)
	Self-efficacy	38.00(8.05)	36.04(8.67)
	Anxiety	8.61(6.02)	11.07(6.31)
Nullified: 3 males & 16 females	Performance	0.83 (0.58)	2.03 (1.57)
	Self-efficacy	30.33(20.31)	30.95(9.40)
	Anxiety	7.00(3.60)	11.31(4.77)
Implicit: 12 males & 22 females	Performance	2.69 (1.40)	2.36 (1.24)
	Self-efficacy	34.00(6.82)	36.68(7.49)
	Anxiety	9.67(2.99)	10.91(5.94)

statutory requirements. All Institutional Review Board protocols were followed including the collection of signed consent forms, voluntary and confidential participation, and debriefing after the study materials were collected.

RESULTS

Analyses of Stereotype Threat Conditions by Gender

A 2x4 MANOVA was used to determine if differences in student performance, test-specific statistics self-efficacy, and test-specific statistics test-anxiety differed by gender (two groups: male and female) and ST condition (four groups: implicit ST, explicit ST, nullified ST, and reverse ST). No significant differences were found for the main effects of gender ($F[3, 91]$, Pillai's trace=.08, $p=.07$) or the interaction effect of gender by condition ($F(9, 279)$, Pillai's trace=.13, $p=.19$). Descriptive statistics are reported in **Table 2**. Six students marked that they were interested in pursuing a college major or career in statistics; five of those students were male.

Analysis of Open-Ended Items

The free-response questions included in the survey provided an opportunity to measure the students' perceptions of the role of gender on mental capacity and opportunities in statistics.

- Do you feel that men and women have the same mental capacity to achieve in statistics? How about the same opportunities?

A vast majority of the participants (91.55%) stated that gender does not influence one's mental capacity, while 8.45% stated that men and women do not have the same mental capacity. When broken down by gender, 67% of the students who stated that men and women do have the same mental capability were women and 5.63% of the students who stated that men and women do not have the same mental capacity were women. A Chi-square test of independent was not significant for gender by yes/no response for mental capacity, $\chi^2(1, 71)=.24$, $p=0.89$. When asked about opportunities in statistics, 49.30% reported that men and women have equal opportunities in statistics and 50.70% reported that men tend to be favored with better opportunities. No one reported that women have more opportunities to achieve in statistics. When broken down by gender, 28.17% of the women stated that men and women have the same opportunities in statistics; 45.07% of the women stated that men have greater opportunities in statistics. A Chi-square test of independent was significant for gender by yes/no response for opportunity, $\chi^2(1, 71)=16.49$, $p<.001$. Fewer women believed that there were equal opportunities for them to achieve in statistics.

DISCUSSION

This is the first study to study ST in statistics. Our results indicated that ST instructions did not impact the students' performance on statistics problems. Thus, it appears that the negative effects of ST on females that are found in mathematics are not an issue with this sample of undergraduate statistics students. We also failed to find differences for self-efficacy and test-anxiety by gender and ST condition.

Null findings are an important part of the research process and important for filling gaps in research because they tell us what is not important or effective. Unfortunately, it is well documented that research with significant results is more likely to be published, which results in a publication bias (Fanelli, 2010). However, this is at the expense of the principles of the scientific method (Campbell et al., 2020). Studies that "show null results despite sufficient statistical power are part of the research process and are fundamental to informing the next research question to be tested" (Campbell et al., 2020, p. 1).

The larger number of women than men enrolled in the elementary statistics course may have attenuated the impact of ST instructions. The problems given to the student may have also been too simple for them. The mean score for the four problems was $M=2.75$, $SD=1.38$. More complicated problems may have allowed for the effect of ST and anxiety to appear. It also may be that motivation and performance are not contributing factors to the gender disparities in statistics participation, but that other factors, such as the perception of statistics being a White and male dominated field (Taasobshirazi et al., 2022), are the contributing factors. Indication of this perception that males have greater opportunities to achieve in statistics were seen in our open-ended items.

Despite the lack of ST effects related to gender in these statistics courses, it is possible that educational and social contexts could still create hostile conditions by which gender differences may be triggered. For example, persistent negative gender stereotyped messages from an instructor may lead to the development of a threat. Our study only looked at a single time point, and the ST instructions were delivered by a researcher with whom students did not have a relationship and who was unrelated to the course.

Although most of the students stated that men and women had the same mental capacity for statistics, a large percentage of students, particularly women, commented that men have more opportunities to succeed in statistics. The most common reason that women reported men having more opportunities in statistics is sexism in the field, specifically a bias against women, stemming from a belief that women

are inferior in science and mathematics (Allen et al., 2022). About 57% of the free-response answers to the question about opportunities cited sexism against women and noted that society doubts their capabilities. About 43% of the responses claimed that men are more desired in the STEM field or that society views women in STEM unconventional.

Much more research is needed to understand the cognitive, motivational, and social variables that may contribute to gender disparities in statistics participation. Longitudinal research with statistics students at different levels of study can help pinpoint when these disparities begin. There is little information regarding when and why gender disparities in statistics achievement, if any, may begin and how they are linked to participation in statistics classes and programs later down the road. Multilevel and structural equation models can help researchers understand mediating, moderating, and contextual factors. Finally, in depth interviews can help researchers better understand students' thoughts about ST conditions, their perceptions of statisticians, and their perceptions of the welcomeness of the field.

Author contributions: All authors were involved in concept, design, collection of data, interpretation, writing, and critically revising the article. All authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: Authors declared that the study was approved by the Kennesaw State University Institutional Review Board (IRB-FY22-67, Approved 8/2021).

Declaration of interest: Authors declare no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

- Allen, D., Dancy, M., Stearns, E., Mickelson, R., & Bottia, M. (2022). Racism, sexism and disconnection: contrasting experiences of Black women in STEM before and after transfer from community college. *International Journal of STEM Education*, 9(1), 1-21. <https://doi.org/10.1186/s40594-022-00334-2>
- Baloglu, M. (2001). *An application of structural equation modeling techniques in the prediction of statistics anxiety among college students* [Unpublished doctoral dissertation]. Texas A&M University-Commerce.
- Batanero, C., Merino, B., & Díaz, C. (2003). Assessing secondary school student's understanding of average. *European Research in Mathematics Education III*, 3, 1-9.
- Boas, L. V. (2020). *Diversity in data science: A systemic inequality how FAANG companies are dealing with this structural problem*. Towards Data Science. <https://towardsdatascience.com/diversity-in-data-science-a-systemic-inequality-b97a0e953f6e>
- Brown, P. L., Concannon, J. P., Marx, D., Donaldson, C. W., & Black, A. (2016). An examination of middle school students' STEM self-efficacy with relation to interest and perceptions of STEM. *Journal of STEM Education*, 17, 27-38.
- Buck, J. L. (1985). A failure to find gender differences in statistics achievement. *Teaching of Psychology*, 12(2), 100-100.
- Campbell, K. L., Moore, J. B., & Bartholomew, J. B. (2020). The importance of publishing null results: Editorial guidelines to contribute to the reduction of publication bias in translational exercise research. *Translational Journal of the American College of Sports Medicine*, 5(11), 1. <https://doi.org/10.1249/TJX.000000000000141>
- Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, 89(2), 318-328. <https://doi.org/10.1037/0022-0663.89.2.318>
- Charles, K. K., & Luoh, M. C. (2003). Gender differences in completed schooling. *Review of Economics and Statistics*, 85(3), 559-577. <https://doi.org/10.1162/003465303322369722>
- Criado-Perez, C. (2019). *Invisible women: Exposing data bias in a World designed for men*. Abrams Press.
- Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985). *Development and validation of an instrument to measure statistical anxiety* [Paper presentation]. The Annual Meeting of the American Statistical Association Statistics Education Section.
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, 28(12), 1615-1628. <https://doi.org/10.1177/014616702237644>
- Es, C. V., & Weaver, M. M. (2018). Race, sex, and their influences on introductory statistics education. *Journal of Statistics Education*, 26, 48-54. <https://doi.org/10.1080/10691898.2018.1434426>
- Evans, T., Thomas, M. O., & Klymchuk, S. (2021). Non-routine problem solving through the lens of self-efficacy. *Higher Education Research & Development*, 40(7), 1403-1420. <https://doi.org/10.1080/07294360.2020.1818061>
- Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLoS ONE*, 5(4), e10068. <https://doi.org/10.1371/journal.pone.0010068>
- 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. <https://doi.org/10.3758/BRM.41.4.1149>
- Fogliati, V. J., & Bussey, K. (2013). Stereotype threat reduces motivation to improve: Effects of stereotype threat and feedback on women's intentions to improve mathematical ability. *Psychology of Women Quarterly*, 37(3), 310-324. <https://doi.org/10.1177/0361684313480045>
- Forbes. (2020). *Fifteen most valuable college majors*. Retrieved January 2, 2023. <https://www.forbes.com/pictures/lmj45jgfi/no-15-statistics/#1448e884a4e31>
- Glassdoor (2020). 50 best jobs in America. https://www.glassdoor.com/List/Best-Jobs-in-America-LST_KQ0,20.htm
- Hiller, S. E., Kitsantas, A., Cheema, J. E., & Poulou, M. (2021). Mathematics anxiety and self-efficacy as predictors of mathematics literacy. *International Journal of Mathematical Education in Science and Technology*, 53(8), 2133-2151. <https://doi.org/10.1080/0020739X.2020.1868589>

- Jacobs, J. E. (2005). Twenty-five years of research on gender and ethnic differences in math and science career choices: What have we learned?, *New Directions for Child and Adolescent Development*, 110, 85-94. <https://doi.org/10.1002/cd.151>
- 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. <https://doi.org/10.1080/00220973.2011.567312>
- Kapitanoff, S., & Pandey, C. (2017). Stereotype threat, anxiety, instructor gender, and underperformance in women. *Active Learning in Higher Education*, 18(3), 213-229. <https://doi.org/10.1177/1469787417715202>
- Legaki, N. Z., Xi, N., Hamari, J., Karpouzis, K., & Assimakopoulos, V. (2020). The effect of challenge-based gamification on learning: An experiment in the context of statistics education. *International journal of human-computer studies*, 144, 102496. <https://doi.org/10.1016/j.ijhcs.2020.102496>
- Mathnasium. (2022). Math vs. statistics: Important points one should know. <https://www.mathnasium.ca/eastregina/news/celebrate-mathematics-and-statistics-awareness-month-this-april>
- Navidi, W. C., & Monk, B. J. (2019). *Elementary statistics*. McGraw-Hill.
- NSF. (2017). Women, minorities, and persons with disabilities in science and engineering: 2017. *National Science Foundation*. www.nsf.gov/statistics/wmpd/
- 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. <https://doi.org/10.1177/0146167203252810>
- Onwuegbuzie, A. J. (1995). Statistics test anxiety and female students. *Psychology of Women Quarterly*, 19, 413-418. <https://doi.org/10.1111/j.1471-6402.1995.tb00083.x>
- Pilotti, M. A. (2021). What lies beneath sustainable education? Predicting and tackling gender differences in STEM academic success. *Sustainability*, 13(4), 1671. <https://doi.org/10.3390/su13041671>
- 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. <https://doi.org/10.1177/0013164493053003024>
- 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. <https://doi.org/10.1111/0022-4537.00201>
- Rossman, A., Chance, B., Medina, E., & Obispo, C. P. S. L. (2006). Some key comparisons between statistics and mathematics, and why teachers should care. In *Thinking and Reasoning with Data and Chance: Sixty-Eighth Annual Yearbook of the National Council of Teachers of Mathematics* (pp. 323-333).
- Saidi, S. S., & Siew, N. M. (2019). Assessing students' understanding of the measures of central tendency and attitude towards statistics in rural secondary schools. *International Electronic Journal of Mathematics Education*, 14(1), 73-86. <https://doi.org/10.12973/iejme/3968>
- Schram, C. M. (1996). A meta-analysis of gender differences in applied statistics achievement. *Journal of Educational and Behavioral Statistics*, 21(1), 55-70. <https://doi.org/10.3102/10769986021001055>
- Smith, J. L., & White, P. H. (2001). Development of the domain identification measure: A tool for investigating stereotype threat effects. *Educational and Psychological Measurement*, 61(6), 1040-1057. <https://doi.org/10.1177/00131640121971635>
- 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. <https://doi.org/10.1023/A:1021051223441>
- 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. <https://doi.org/10.1006/jesp.1998.1373>
- 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. <https://doi.org/10.1037/0022-3514.69.5.797>
- Stern, D., Stern, R., Parsons, D., Musyoka, J., Torgbor, F., & Mbasu, Z. (2020). Envisioning change in the statistics-education climate. *Statistics Education Research Journal*, 19(1), 206-225. <https://doi.org/10.52041/serj.v19i1.131>
- Stroup, D. F., & Jordan, E. W. (1982). Statistics: Monster in the university. In *Proceedings of Statistical Education, the American Statistical Association* (pp. 135-138).
- Susbiyanto, S., Kurniawan, D. A., Perdana, R., & Riantoni, C. (2019). Identifying the mastery of research statistical concept by using problem-based learning. *International Journal of Evaluation and Research in Education*, 8(3), 461-469. <https://doi.org/10.11591/ijere.v8i3.20252>
- Taasobshirazi, G., Wagner, M., Brown, A., & Copeland, C. (2022). An evaluation of college students' perceptions of statisticians. *Journal of Statistics and Data Science Education*, 30(2), 138-153. <https://doi.org/10.1080/26939169.2022.2058655>
- Witherspoon, E. B., & Schunn, C. D. (2020). Locating and understanding the largest gender differences in pathways to science degrees. *Science Education*, 104, 144-163. <https://doi.org/10.1002/sce.21557>
- Woehlke, P. L., & Leitner, D. W. (1980). Gender differences in performance on variables related to achievement in graduate-level educational statistics. *Psychological Reports*, 47(3), 1119-1125. <https://doi.org/10.2466/pr0.1980.47.3f.1119>
- Yates, A., Starkey, L., Egerton, B., & Flueggen, F. (2021). High school students' experience of online learning during COVID-19: The influence of technology and pedagogy. *Technology, Pedagogy and Education*, 30(1), 59-73. <https://doi.org/10.1080/1475939X.2020.1854337>