The Contributions of Self-Efficacy and Test Anxiety to Physics Exam "Gender Gaps"

by Jared B. Stang, Emily Altiere, Patrick J. Dubois, Yulia Egorova, Christine Goedhart, Joss Ives, Karen M. Smith and Jaclyn J. Stewart

Jared B. Stang <jared.stang@ ucalgary.ca> Department of Physics and Astronomy, University of Calgary, Calgary, Alberta

Women are underrepresented relative to men in science, technology, engineering, and mathematics (STEM) in Canada [1]. The demographic make-up of post-secondary STEM students skews toward men — particularly in physics [2] — and there was also found a gap in the retention of first-year women in STEM programs from 2010–2015, with 66% of women persisting compared to 72% of men [1]. These imbalances are unjust and they ultimately reduce the quality of physics work in Canada [3].

The question of what contributes to this gender imbalance is complex, however, symptoms of inequities can be found. Prior to the study reported here, author J.B.S. compiled exam score and gender data across four introductory physics courses at the University of British Columbia (UBC), comprising more than 2500 engineering, science, and non-science students. There was a consistent difference — between 2% and 7% — in the exam scores in favour of men. This type of "gender gap" in assessment scores in has also been reported for research-validated concept inventories at Canadian universities [4,5] and in the wider Physics Education Research community [6]. The widespread nature of the gaps in assessment scores provides evidence that gender-based inequities permeate physics education. Admission to undergraduate programs, scholarships, and obtainment of research experience depends on grades, and these inequities in grading may exclude more women than men. However, a focus on exam scores cannot explain why the differences exist or give any suggestions to instructors looking to counteract this phenomenon. In this article, we explore two psychological factors which may contribute to gender-based differences in exam scores.

Summary

We show that in our active-learning introductory physics course both physics self-efficacy and test anxiety may contribute to gender-based disparities in assessment scores.

Self-efficacy is the belief in one's capabilities to complete a particular task. This is an important construct because it influences the choices a person makes and the effort they will put into a task [7]. Recent reviews show that in general, women enter STEM programs with lower self-efficacy [8], and in particular report lower levels of physics self-efficacy than men [9]. Self-efficacy has also been shown to be related to performance [10,11] and persistence [12] in physics. Test anxiety is the tendency to become anxious in a high-stakes testing situation. If one group overall experiences higher levels of test anxiety, this may impair the performance of that group on exams, creating a disadvantage. Test anxiety has been shown to mediate performance for women — but not men — in an introductory biology course [13] and, more recently, across science courses [14].

We examine the physics self-efficacy and test anxiety reported by students in our introductory algebra-based physics course, and the relationship of these to exam scores. The study shared here is a preliminary step toward better understanding more aspects of the student experience in our physics course. The wider aim of this program of research is to identify the relevant affective factors, and their causes, to make concrete suggestions about how to make physics education more inclusive.

MEASURING STUDENT PERFORMANCE, SELF-EFFICACY, AND TEST ANXIETY

This study took place in a grade-12 equivalent introductory physics course at the UBC Vancouver campus. The course was taught using active-learning techniques, including personal response "clicker" questions and worksheet activities. The approximately 800 students were enrolled across three lecture sections, each taught by a white male instructor.

Even though less than 1% of students taking this course go on to do a physics degree, the results from this study are important to the goal of increasing the participation of women in physics for several reasons. First, these students are members of the wider civic community and their experiences in physics and science will impact how they participate now and in the future. Second, addressing inequities across all the courses we teach is an important way to help shift the culture of physics towards more inclusion. Finally, the trends shown here may be applicable to courses for physics majors, helping us to understand why students who begin in the field do not persist.

For this study, we used course performance data, data from an administered diagnostic test and affective and demographic data collected through surveys. The design was informed by a pilot study whose results are presented in Ref. 15. A total of 622 unique participants responded to our demographic survey. Although the survey contained many items, we focus here on our gender identity question, for which 170 students selected "Man" and 444 students selected "Woman". Because of the insufficient sample size, students who self-identified as members of other gender identity groups were excluded from the analysis presented here. We aspire to incorporate the experience of those identifying as non-binary persons in future work. To measure the incoming physics knowledge of our students, we used the Force Concept Inventory, a well-established diagnostic survey measuring the degree to which a person has a Newtonian view of mechanics concepts. To assess students' self-efficacy and test anxiety, we administered an affective survey with questions derived originally as a subset of the Motivated Strategies for Learning Questionnaire (MSLQ) [16]. The survey items were presented using a 7-point scale with options ranging from "strongly disagree" to "strongly agree". Responses to each item were translated to a numerical scale, and the items within each construct were averaged to get values for each student's self-efficacy and test anxiety. Performance data came in the form of grades from the final exam, which consisted of 15 multiple choice questions and 5 multi-part open-response problems.

Our multiple linear regression model for predicting final exam score was chosen in advance of analyzing these data, and used gender identity, self-efficacy, test anxiety, and FCI score as predictor variables. For the regression model, all variables were standardized to have means of zero and standard deviations of one.

SELF-EFFICACY AND TEST ANXIETY DISPARITIES AND THEIR RELATION TO EXAM SCORES

Women had lower self-efficacy, higher test anxiety, and lower FCI scores, while the final exam scores show no differences by gender identity (see Fig. 1). Although the exam scores did not show a significant difference this year — in contrast to both the pilot study and the informal data collection that inspired this study — building a regression model allows us to see how self-efficacy and test anxiety are related to exam scores, which may help us understand the gender-based differences that typically appear.

The model to predict the final exam score, with the numerical regression coefficients, looks as follows,

$$\text{final exam} \sim 0.29 \times \text{gender identity} + 0.26 \times \text{self-efficacy} - 0.14 \times \text{test anxiety} + 0.33 \times \text{pre-FCI},$$

where all four regression coefficients are significant at or beyond the $p < .01$ level. The fraction of the total variance in the final exam scores explained by this model is $R^2 = .32$.

This model shows that higher self-efficacy is associated with a higher final exam score while higher test anxiety is associated with a lower final exam score. The coefficient for self-efficacy in the linear model is larger than that for test anxiety and the mean difference between women and men in self-efficacy is larger than the mean difference in test anxiety. This suggests that self-efficacy may play a larger role in gender-based disparities on assessments than test anxiety. This is consistent with the pilot study, where test anxiety did not improve the explanatory power of the models considered. While there was no raw gender-based difference in mean exam scores, the model predicts that for students with identical self-efficacy, test anxiety, and FCI score, women will perform better than men on the final exam by 0.29 standard deviations (or about 5%). This is an encouraging trend which suggests that, in this course, initial performance disparities as measured by the FCI are either not appearing in or tending to even out by the final exam.

The predictive value of physics self-efficacy suggests that attending to self-efficacy through facilitation and curriculum has the potential to improve equity in the classroom. Kalendar *et al.* [11] summarize some approaches that an instructor might consider, such as: making sure that equity and inclusion are part of the course design, including explicitly valuing contributions from all students; assigning roles in group work to reduce unbalanced participation; and fostering a growth mindset so that students view challenges as learning opportunities. Earlier work by Sawtelle [17] identifies that certain classroom components — including cooperative group work, instruction focused on physics model development and use, and instructor interaction — are mechanisms through which the physics self-efficacy of students may be positively impacted. However, in their recent review, Henderson *et al.* [9] point out that there is a deficit of evidence describing classrooms which successfully bolster the self-efficacy of students (and women in particular). Given the demonstrated potential for self-efficacy to improve the learning and performance of all students, there is a need for teaching innovation and further research focused on this aspect of the student experience.

Test anxiety had a negative impact on exam scores, and women reported a higher level of test anxiety than men in our study, suggesting that strategies to mitigate test anxiety may be one path toward a more equitable classroom. Ballen *et al.* [13] and Salehi *et al.* [14] are in agreement that rethinking

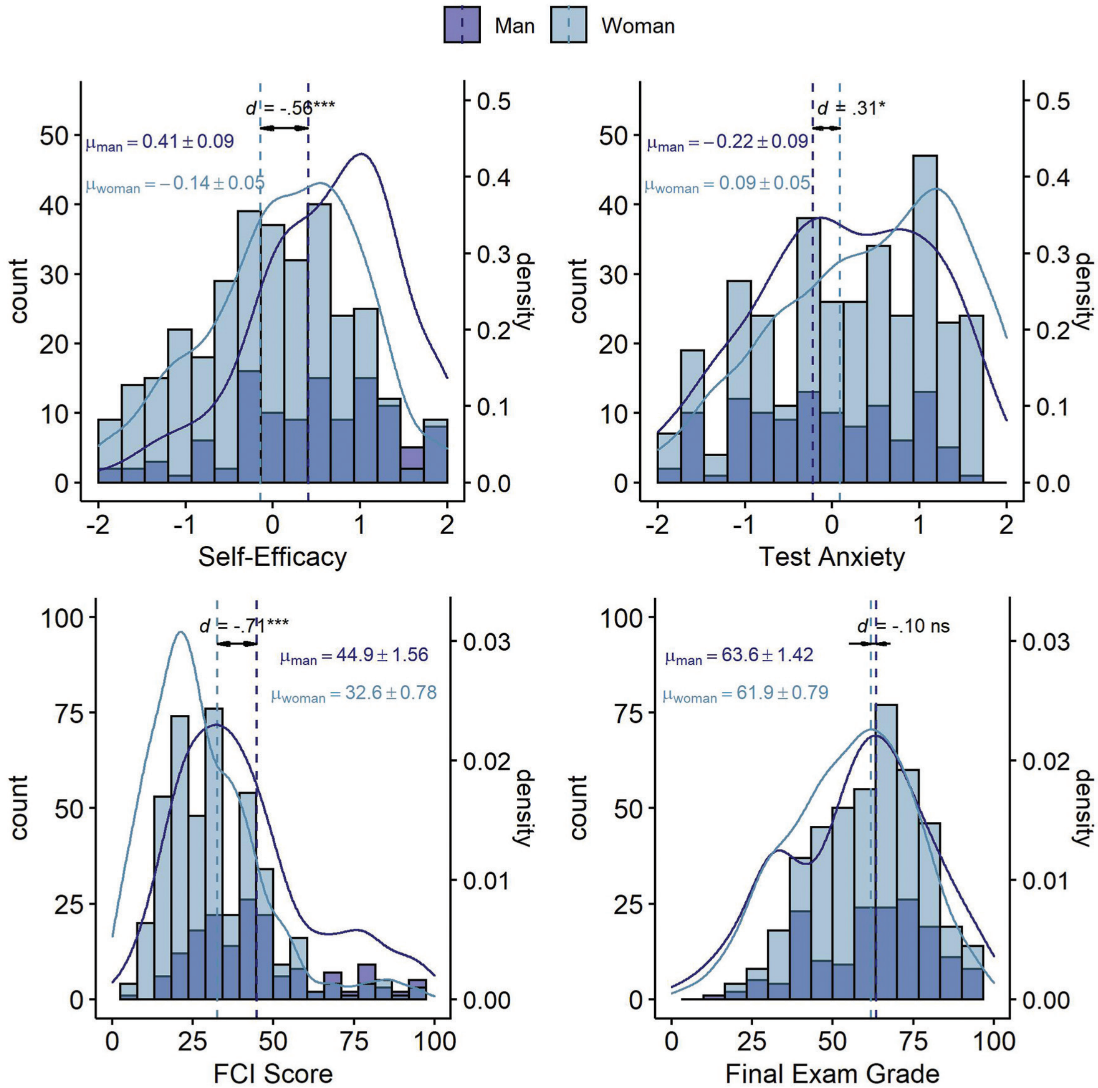

Fig. 1. Histograms and overlaid density plots comparing the distributions of women and men on self-efficacy, test anxiety, incoming FCI score, and final exam grade. Self-efficacy and test anxiety have been standardized while the FCI and final exam grades are on a percentage scale. Means and standard error in the means are indicated for the two gender identity groups on each panel, and the effect size Cohen's d for the difference between men and women is shown. Statistical significance is evaluated using a two-tailed t-test assuming unequal variance and using an $n = 5$ Bonferroni correction for multiple comparisons (including a comparison of end-of-term FCI scores not presented here). ns: not significant; *: $p < .01$; **: $p < .002$; ***: $p < .0002$.

assessment strategies — including devaluing high-stakes exams in favour of multiple lower-stakes and mixed assessment types — is the most promising approach to reducing the possible negative and biased impact of test anxiety. Harris *et al.* [18] recently reported that a combination of interventions designed to reduce test anxiety did not reduce reported levels of test anxiety, but did improve exam scores for all students. Unfortunately, in their study, the gender gap in exam scores persisted. As for self-efficacy, more work is needed to better understand what spaces and practices allow students to learn and demonstrate their mastery without the negative effects of test anxiety.

CONCLUSION

This study provides evidence that addressing self-efficacy and test anxiety disparities may be mechanisms to promote equity in the physics classroom. The overall effect of test anxiety on exam scores was smaller than that of self-efficacy in both our pilot study and the data presented here. However, as Madsen *et al.* [6] describe, "the gender gap is most likely due to the combination of many small factors rather than any one factor that can easily be modified." Therefore, all factors should be considered in the pursuit of an equitable classroom environment. In future work, we will include other constructs which have been shown to be important for student achievement and persistence in STEM, such as sense of belonging and science identity [19].

We have presented here data relevant to the wider issue of participation inequities in physics in Canada. Since exam and course scores impact the opportunities available to students and the future choices they make, gender-based disparities in these scores very likely contribute to the overall participation differences. Better understanding how students experience our classrooms by attending to affective dimensions such as self-efficacy and test anxiety can reveal mechanisms responsible for the inequities we observe and, ultimately, help to create classroom environments that are inclusive for all students.

ACKNOWLEDGEMENTS

We acknowledge support from the UBC Equity and Inclusion Scholars Program, the UBC Centre for the Integration of Research, Teaching and Learning, and the UBC Institute for the Scholarship of Teaching and Learning. We thank the students and instructors of the course for their participation.

REFERENCES

1. K. Wall, “Persistence and representation of women in STEM programs”, *Insights on Canadian Society*, Statistics Canada Catalogue no. 75-006-X, ISSN 2291-0840 (2019).
2. A. Perreault, L. Shannon, A. Farenhorst, C. Mavriplis, E. Langelier, and T. Franz-Odendaal, “Analysis of the distribution of gender in STEM fields in Canada”, NSERC Chairs for Women in Science and Engineering report (2018), <http://wiseatlantic.ca/wp-content/uploads/2018/03/WISEReport2017_final.pdf>.
3. E. Haine-Bennett, H.B. Bergsieker, I.R. Coe, A. Koch-Kraft, E. Langelier, S. Morrison, K. Nikoleyczik, T. Schmader, O. Trivailo, S. Twine, and J.E. Decker, “Enacting workplace culture change for excellence in research: a gender lens”, *FACETS*, **5**, 228-233 (2020).
4. M. Normandeau, S. Iyengar, and B. Newling, “The Presence of Gender Disparity on the Force Concept Inventory in a Sample of Canadian Undergraduate Students”, *Canadian Journal for the Scholarship of Teaching and Learning*, **8**, 9 (2017).
5. J. Day, J.B. Stang, N.G. Holmes, D. Kumar, and D.A. Bonn, “Gender gaps and gendered action in a first-year physics laboratory”, *Physical Review Physics Education Research*, **12**, 020104 (2016).
6. A. Madsen, S.B. McKagan, and E.C. Sayre, “Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?”, *Physical Review Special Topics-Physics Education Research*, **9**, 020121 (2013).
7. A. Bandura, “Self-efficacy: toward a unifying theory of behavioral change”, *Psychological Review*, **84**, 191 (1977).
8. S.L. Eddy and S.E. Brownell, “Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines”, *Physical Review Physics Education Research,* **12**, 020106 (2016).
9. R. Henderson, V. Sawtelle, and J.M. Nissen, “Gender and Self-Efficacy: A Call to Physics Educators”, *The Physics Teacher*, **58**, 345-348 (2020).
10. E.M. Marshman, Z.Y. Kalender, T. Nokes-Malach, C. Schunn, and C. Singh, “Female students with A’s have similar physics self-efficacy as male students with C’s in introductory courses: A cause for alarm?”, *Physical Review Physics Education Research*, **14**, 020123 (2018).
11. Z.Y. Kalender, E. Marshman, C.D. Schunn, T.J. Nokes-Malach, and C. Singh, “Damage caused by women’s lower self-efficacy on physics learning”, *Physical Review Physics Education Research*, **16**, 010118 (2020).
12. V. Sawtelle, E. Brewe, and L.H. Kramer, “Exploring the relationship between self-efficacy and retention in introductory physics”, *Journal of research in science teaching*, **49**, 1096-1121 (2012).
13. C.J. Ballen, S. Salehi, and S. Cotner, “Exams disadvantage women in introductory biology”, *PloS one*, **12**, e0186419 (2017).
14. S. Salehi, S. Cotner, S.M. Azarin, E.E. Carlson, M. Driessen, V.E. Ferry, W. Harcombe, S. McGaugh, D. Wassenberg, A. Yonas, and C.J. Ballen, “Gender Performance Gaps Across Different Assessment Methods and the Underlying Mechanisms: The Case of Incoming Preparation and Test Anxiety”, *Frontiers in Education*, **4**, 107 (2019).
15. J.B. Stang, E. Altiere, J. Ives, and P. Dubois, “Exploring the contributions of self-efficacy and test anxiety to gender differences in assessments”, paper presented at the *Physics Education Research Conference 2020* (2020), <https://arxiv.org/abs/2007.07947>.
16. P. Pintrich, D. Smith, T. Garcia, and W. McKeachie, “A manual for use of the Motivated Strategies for Learning Questionnaire”, Technical report no. 91-B-004, The Regents of the University of Michigan (1991).
17. V. Sawtelle, “A Gender Study Investigating Physics Self-Efficacy”, *FIU Electronic Theses and Dissertations*, 512 (2011), <https://digitalcommons.fiu.edu/etd/512>.
18. R.B. Harris, D.Z. Grunspan, M.A. Pelch, G. Fernandes, G. Ramirez, and S. Freeman, “Can Test Anxiety Interventions Alleviate a Gender Gap in an Undergraduate STEM Course?”, *CBE — Life Sciences Education*, **18**, ar35 (2019).
19. G. Trujillo and K.D. Tanner, “Considering the role of affect in learning: Monitoring students’ self-efficacy, sense of belonging, and science identity”, *CBE — Life Sciences Education*, **13**, 6-15 (2014).