Gender differences in self-efficacy states in high school physics

Jayson M. Nissen

Department of Science Education, California State University, Chico, California 95929, USA

(Received 23 October 2018; published 5 March 2019)

Self-efficacy, the belief in one’s ability to succeed in learning tasks, predicts learning and success in education broadly and physics specifically. While self-efficacy increases for students in most introductory science and mathematics courses, self-efficacy consistently decreases for women in physics courses. This study used the experience sampling method to investigate gender differences in the self-efficacy states high school students experienced in physics, other math and science classes, and other classes. Data for the study came from the Sloan survey of youth and social development and included data from 1332 students at 12 different schools collected between 1993 and 1997. Principle components analysis identified a self-efficacy state measure within the data. Comparisons of self-efficacy states across gender and activity (e.g., physics and science and mathematics courses) using a two-level hierarchical linear model identified a large gender difference in self-efficacy states in physics and only in physics. These results add to the growing evidence that female students’ physics self-efficacy tends to decrease after taking physics courses. Given that self-efficacy predicts career choice and success, decreases in women’s physics self-efficacy may inform why women are much less likely to pursue physics careers than men.

DOI: 10.1103/PhysRevPhysEducRes.15.013102

1. INTRODUCTION

Self-efficacy is the belief in one’s ability to succeed in a given task or domain [1]. In university physics courses, self-efficacy predicts students’ achievement [2,3]. For undergraduate physics researchers, shifts in knowledge of science couple with shifts in physics self-efficacy [4]. These relationships in physics align with the broader findings on self-efficacy being related to important academic and career outcomes. Increased self-efficacy causes improved cognitive performance [5], increased interest in pursuing science degrees [6], and improved academic outcomes [7]. Self-efficacy also predicts students taking harder courses [8], academic success [9–11], the college major students choose [8,12], and students’ choice of career [13–15].

Two recent reviews [16,17] cover several gender differences across science, technology, engineering, and mathematics (STEM) disciplines and found that gender differences in self-efficacy exist across most of the STEM domains. Cheryan et al. [17] found that gender differences in self-efficacy, masculine cultures, and a lack of early educational experiences in the disciplines explained the lower rates of participation for women in physics, computer science, and engineering than in biology, chemistry, and mathematics. Eddy and Brownell [16] (p. 10) point out a major limitation of the articles they reviewed, “the majority of [the studies] only sampled students once. Without following students through time, it is impossible to determine the impact of a degree program or course.”

Given the relationships between self-efficacy and student learning, career choice, and persistence, a reasonable goal for physics education is to support students in developing self-efficacy. I reviewed the self-efficacy literature to find studies that used pretests and post-tests and disaggregated the data by gender. I reasoned that such measures infer a relationship between taking a course and shifts in self-efficacy. Students’ self-efficacy tends to decrease [2,18–21] or at best not change [18,19,22] from pre- to postinstruction in introductory physics courses. Female students’ self-efficacy generally decreases more than for their male peers [2,18–21]. In some physics courses, male students’ self-efficacy slightly increased [19,22]. In contrast, students’ self-efficacy increased in introductory chemistry courses [23–25], an introductory algebra course [26], and introductory biology courses [27–29]. These studies in chemistry, biology, and mathematics did not disaggregate the data by gender. However, the increase for all students likely represented an increase for women due to their high level of representation in the data [23–27,29]. I did not identify any studies in engineering or computer science that measured self-efficacy with both pretests and post-tests.

The consistent negative shift in female students’ self-efficacy in physics compared to the less negative shift for
male students’ in physics and the positive shifts in other science and math domains indicates that college physics instruction may harm female students’ physics self-efficacy. Since these studies on self-efficacy in college STEM courses have only focused on one domain and have used pre- and postsurveys to measure self-efficacy, they have not established physics instruction as causing the harm to female students’ physics self-efficacy. For example, these studies leave open the possibility that male and female students have similar experiences in physics but reflect differently on these experiences, known as a state-trait discrepancy. They also leave open the possibility that the students in physics differ from the students in other STEM courses.

To address these limitations and investigate if physics instruction harms female students’ physics self-efficacy, Nissen and Shemwell [21] differentiated between two types of self-efficacy: states, which are dynamic and momentary, and traits, which are more stable attitudes. They characterized the common method of measuring students’ self-efficacy using pre- and postsurveys as trait measures because of their retrospective nature. They used an in-the-moment sampling technique, the experience sampling method (ESM), to measure college students’ self-efficacy states throughout their daily lives. Their results showed that female students experienced much lower self-efficacy states in physics than male students did and that female students’ self-efficacy states in physics were amongst the lowest self-efficacy states they experienced in any activity. In contrast, male and female students had similar self-efficacy states in other STEM courses, and those states were similar to male students’ self-efficacy states in physics. Given that pre- and post-surveys found that female students had larger decreases in their self-efficacy traits than their male peers and the large gender difference in self-efficacy states in physics, Nissen and Shemwell concluded the physics course they investigated harmed female students’ self-efficacy. Their findings and the consistent decreases in female students physics self-efficacy reported in the literature indicate the need for further investigations of self-efficacy in physics to understand the size, extent, and causes of gender differences in physics self-efficacy.

II. PURPOSE

My purpose in this study was to investigate the extent to which female high school students experienced lower self-efficacy states while learning physics than their male peers. To support interpreting gender differences in physics, I also investigated gender differences in self-efficacy states in other school and nonschool activities. Given the consistent gender differences in college physics self-efficacy, I reasoned that gender differences in high school physics would indicate that gender differences in self-efficacy commonly occur in physics.

III. THE EXPERIENCE SAMPLING METHOD DATA SET

The data for the investigation came from a publicly available ESM dataset collected in the 1990s as a part of the Sloan study of youth and social development [30] (SSYSD). The SSYSD collected ESM data in three waves (in 1993, 1995, and 1997) from 4816 students at 33 schools in 12 locations selected to represent the United States. Students participating in the ESM were semirandomly signaled approximately every two hours throughout their day, from 7:30 a.m. to 10:30 p.m., for a one-week period. The signal directed them to complete a one-page survey (available in Ref. [31]) that consisted of open-ended questions about what they were doing and Likert-scale affective questions about their experience. They were instructed to complete the survey as soon after being signaled as possible. Csikszentmihalyi and Schneider [32] describe the methods used in the SSYSD ESM data collection. The data slightly overrepresented female students’ experiences and slightly underrepresented experiences in the evening, but was otherwise representative of adolescents’ experiences [33].

The age of the SSYSD data leaves open the possibility it no longer represents students’ experiences. During the intervening time between this study and the data collection, many initiatives, such as the Next Generation Science Standards, have sought to improve high school science instruction. However, these initiatives have not consistently or meaningfully changed gender differences in physics: the percentage of women undergraduate physics degree recipients remained steady over the last twenty years [34] as has the percentage of students that take high school physics who are female [35]. Given the static representation of women in physics, and the correlations between self-efficacy and both choice of major and choice of courses [8,12], the gender differences in the self-efficacy states that students experienced in high school physics has likely not changed in the last 25 years. The age of the SSYSD data also has the advantage of informing the extent and durability of any gender differences in physics self-efficacy. Consistent gender differences that have persisted through many initiatives to improve high school and college physics instruction will require new approaches to investigate and address than differences that occur only recently and only under some circumstances.

The SSYSD data were not designed to measure students’ self-efficacy states. However, the data likely included a measure of students’ self-efficacy states because the survey included similar questions to the three questions about skill, control, and success that formed the core of self-efficacy state measures in prior ESM research [21,36]. These questions are central to measuring self-efficacy because control and capability are central attributes of self-efficacy [1]. The questions were:
IV. METHODS

I used principle components analysis (PCA) to identify the self-efficacy state construct within the entire SSYSD ESM dataset by checking if the skill, control, and success questions loaded on the same component and identifying other questions that loaded on that component. I used all questions that loaded on the component with skill, control, and success to form the self-efficacy state construct. PCA identifies coherent subsets of variables that are relatively independent [37] and is commonly used in ESM studies [31]. I have included the Cronbach’s $\alpha$ as an additional measure of internal consistency. Once PCA identified the questions that formed the self-efficacy state measure, I averaged the questions on a 5-point scale, 0 (none) to 4 (very much), to form the self-efficacy state measure.

To analyze the gender differences across the school activities while addressing the nested nature of the data, I built a two-level hierarchical linear model in the HLM 7 software. The self-efficacy states (level 1) nested within the students (level 2). I did not include school or class as a third level because the data set did not include this information. The outcome variable for these models was the self-efficacy state measure. The level 1 input variables were physics, math and science classes other than physics, and school. The level 2 input variable was gender, coded as 0 for male and 1 for female. Table I describes the number of students and samples in each category. Equations (1)–(5) show the level-1 and level-2 equations. This structure allowed comparing male and female students' self-efficacy states in physics and provided the gender differences in self-efficacy states in similar (math and science and school) and different (nonschool) activities to support interpreting the differences in physics. The structure also meant the intercept predicted mean self-efficacy states for male students in nonschool activities. I used the hypothesis testing feature of the HLM software to calculate the means and standard error for male and female students in each activity.

**Level-1 Equation**

$$SE_{ij} = \beta_{0j} + \beta_{1j} \times \text{Physics}_{ij} + \beta_{2j} \times \text{math and science}_{ij} + \beta_{3j} \times \text{School}_{ij} + r_{ij}. \quad (1)$$

**Level-2 Equations**

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times \text{Gender}_j + u_{0j}, \quad (2)$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11} \times \text{Gender}_j + u_{1j}, \quad (3)$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21} \times \text{Gender}_j + u_{2j}, \quad (4)$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31} \times \text{Gender}_j + u_{3j}. \quad (5)$$

Physics, math and science, and school were dummy variables coded as either a 0 or 1, were exclusive of each other, and represented students’ response to the main thing they were doing. The SSYSD ESM data set included numeric codes for the main thing that students were doing when they responded, and the coding book is available in Ref. [30]. “Math and science” included mathematics, general science, biology, chemistry, earth science, and computer science. I aggregated the math and science courses to control for a large number of variables decreasing the statistical power of the study and resulting in a false negative result. Prior to aggregating the math and science courses, I analyzed them using $t$ tests and found no meaningful and reliable gender differences, either $d < 0.2$ or $p > 0.05$. “School” included any other activities coded as school. Table I details the distribution of students and experiences across the four activities.

A histogram of male and female student’s self-efficacy states in physics facilitated comparison of the distributions of the two groups experiences and informed the meaningfulness of those differences. The histogram did not take into account the nested nature of the data. Nonetheless, it presented the general range of self-efficacy states that male and female students experienced in high school physics.

### Table I. Structure of the SSYSD ESM data.

<table>
<thead>
<tr>
<th></th>
<th>Physics</th>
<th>Math and Science</th>
<th>School</th>
<th>Nonschool</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Students</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>636</td>
<td>1278</td>
<td>1321</td>
<td>1332</td>
</tr>
<tr>
<td>Male</td>
<td>30</td>
<td>276</td>
<td>568</td>
<td>586</td>
<td>595</td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>360</td>
<td>710</td>
<td>735</td>
<td>737</td>
</tr>
<tr>
<td><strong>Responses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>1338</td>
<td>13181</td>
<td>30156</td>
<td>44745</td>
</tr>
<tr>
<td>Male</td>
<td>48</td>
<td>554</td>
<td>5161</td>
<td>11846</td>
<td>17609</td>
</tr>
<tr>
<td>Female</td>
<td>22</td>
<td>784</td>
<td>8020</td>
<td>18310</td>
<td>27136</td>
</tr>
</tbody>
</table>
V. RESULTS

PCA of the survey responses in the SSYSD data identified a reliable self-efficacy state measure with an eigenvalue of 3.3 that explained 12.2% of the variance in the data. The self-efficacy state measure identified by the PCA included the three core questions on skill, control, and success and three additional questions. Of the three core questions, success had an excellent loading (0.70), control had a very good loading (0.67), and skill had a very good loading (0.62). The strength of these loadings for these core questions supported interpreting this factor as a measure of self-efficacy states. The three other questions that loaded on the self-efficacy state measure were: “Were you living up to your expectations?” (0.76), “Were you living up to expectations of others?” (0.67), and “Did you feel good about yourself?” (0.58). These three additional questions all aligned with the self-efficacy state being a measure of students’ feelings of ability and control in the activity at hand. Cronbach’s $\alpha$ for these 6 questions was 0.71. The strength of the loadings of three core questions; the consistency of the additional three questions with self-efficacy; and the eigenvalue, variance explained, and alpha of these 6 questions indicated that they formed a reliable measure of self-efficacy states.

The hierarchical linear model indicated that female students experienced much lower self-efficacy states in physics than their male peers, $-0.49$ points on a 0 to 4 scale, and this difference was statistically significant ($p = 0.001$), as shown in Table II. No other gender differences were meaningful or statistically significant. The gender difference in physics did not occur because male students had elevated self-efficacy states in physics. Rather, male students’ self-efficacy states were lower in physics than in other math and science courses. As shown in Fig. 1, female students experienced self-efficacy states 0.63 points lower in physics than in other math and science courses. The histograms of self-efficacy states in high school physics, Fig. 2, indicated that female students experienced very low and moderately low self-efficacy states in physics twice as often as male students. These results indicated that the gender difference in self-efficacy states in physics was large, unique, and reliable.

VI. DISCUSSION

The PCA identified a reliable and consistent measure of self-efficacy states. Analysis of the self-efficacy states revealed large gender differences occurred in physics but did not occur in the other activities. The gender differences represented low self-efficacy state experiences for female students in physics compared to other activities, as opposed to high self-efficacy states for male students. Breaking down the self-efficacy states in physics showed female students experienced very low self-efficacy states far more frequently and experienced very high self-efficacy states far less often than their male peers.

While the sample of experiences in physics in this study came from only 14 female students and 30 male students, the results indicated a large and statistically significant difference in boys and girls experiences in physics, and only in physics, that aligns with the gender difference in self-efficacy states in a college physics course [21]. By measuring self-efficacy states, these studies locate the gender difference in self-efficacy within the experience.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\gamma$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>For intercept $\beta_{0j}$</td>
<td>Intercept $\gamma_{00}$</td>
<td>2.93</td>
</tr>
<tr>
<td></td>
<td>Gender $\gamma_{01}$</td>
<td>0.01</td>
</tr>
<tr>
<td>For physics $\beta_{1j}$</td>
<td>Intercept $\gamma_{10}$</td>
<td>$-0.06$</td>
</tr>
<tr>
<td></td>
<td>Gender $\gamma_{11}$</td>
<td>$-0.49$</td>
</tr>
<tr>
<td>For math and science $\beta_{2j}$</td>
<td>Intercept $\gamma_{20}$</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Gender $\gamma_{21}$</td>
<td>$-0.04$</td>
</tr>
<tr>
<td>For school $\beta_{3j}$</td>
<td>Intercept $\gamma_{30}$</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Gender $\gamma_{31}$</td>
<td>0.01</td>
</tr>
</tbody>
</table>
of learning physics. Gender differences in physics self-efficacy states align with prior research indicating that women’s self-efficacy traits tend to decrease after taking physics courses but not after taking other courses and that men’s self-efficacy traits do not tend to decrease after taking physics courses. By measuring self-efficacy traits before and after instruction, these prior studies not only indicate that gender differences in physics self-efficacy traits exist, but show those gender differences tend to increase after students take physics courses due to decreases in women’s physics self-efficacy traits. These results show a reliable gender difference in physics self-efficacy (states or traits) linked to the experience of learning physics: either through the depressed self-efficacy states women experience while learning physics or decreases in women’s physics self-efficacy traits after taking a physics course. The results in this study expand upon the prior work to show that gender differences in physics self-efficacy states have occurred since the 1990s and also occur in high school physics courses.

VII. CONCLUSION

Cheryan et al. [17] found that masculine cultures in physics, engineering, and computer science act to deny women a sense of belonging in those disciplines. That finding aligns with the literature showing decreases in women’s self-efficacy in physics courses but not in mathematics, biology, or chemistry courses and with this and a prior study showing women and girls experience lower self-efficacy states in physics but not in other science and mathematics courses. Women or any group of students who regularly experience very low self-efficacy states in physics courses will not produce many aspiring physicists. While college physics courses cannot change the childhood experiences of women, education researchers can design courses that account for and support the needs of all students. To provide knowledge to educators and instructors to design courses that support students from under-represented groups in physics, research needs to extend beyond the characteristics students bring into the physics course and include aspects of the physics curriculum the instructors and departments control: the teaching methods, pathways into and through physics degrees, and the physical and social environments in which students learn.

One direction for future research is stereotype threat, “a situation where one faces judgment based on societal stereotypes about one’s group” [38] (p. 5) because it may depress the self-efficacy states of women in physics. Consistent with this study’s findings, stereotype threat occurs for women in physics but not in chemistry [39–41]. Stereotype threat induces a stress response [42,43] that undermines learning [44] and performance [45,46], and higher stress correlates with lower self-efficacy [21]. Environmental characteristics such as stereotypical objects in the learning environment [47] and underrepresentation [39,46,48], and student characteristics such as fixed mindsets [49] and stereotype endorsement [50] can all induce stereotype threat and are frequently present in physics courses. For example, Leslie et al. [51] found that physicists view a fixed mindset of raw ability as important to success in physics and this view correlated with the low representation of women in physics compared to other disciplines. These connections are inferential, but they point to research on the links between self-efficacy and stereotype threat as a path toward understanding and improving the self-efficacy states women experience in physics.


