Statistics Related Self-Efficacy: A Confirmatory Factor Analysis Demonstrating a Significant Link to Prior Mathematics Experiences for Graduate Level Students

Karen Larwin*

Educational Foundations Research Technology and Leadership Youngstown State University, United States

Abstract

The present study examined students' statistics-related self-efficacy, as measured with the current statistics self-efficacy (CSSE) inventory developed by Finney and Schraw (2003). Structural equation modeling was used to check the confirmatory factor analysis of the one-dimensional factor of CSSE. Once confirmed, this factor was used to test whether a significant link to prior mathematics experiences exists. Additionally, a new post-structural equation modeling (SEM) application was employed to compute error-free latent variable score for CSSE in an effort to examine the ancillary effects of gender, age, ethnicity, department, degree level, hours completed, expected course grade, number of college-level math classes, current GPA on students' CSSE scores. Results support the one-dimensional construct and as expected, the model demonstrated a significant link between CSSE scores and prior mathematics experiences to CSSE. Additionally the students' department, expected grade, and number of prior math classes were found to have a significant effect on student's CSSE scores.

Keywords: Statistics related self-efficacy, confirmatory factor analysis, error–free latent variable analysis.

1 Introduction

For more than fifty years, researchers have examined the factors that influence student success in statistics coursework. These factors include variables such as prior coursework, gender, age, area of major as well as psychological constructs like self-efficacy regard to statistics. Self-efficacy, specifically perceived self-efficacy, is a social cognitive construct reflecting the degree to which an individual believes he or she can successfully perform a specific behavior or task (Bandura, 1986, 1997) [7,8]. According to Bandura, self-efficacy beliefs can influence a person's choice of tasks, level of effort, persistence, resilience, and ultimately success and achievement. In addition, self-efficacy is situation specific; it is only predictive of the specific behaviors one believes he/she can perform in a specific context; it is not a...
generalized tendency or trait. Bandura (1997) maintains that students appraise their self-efficacy of their performances, past experiences, vicarious experiences, their psychological reactions to experiences and through feedback from others [8]. While many researchers have demonstrated that there is a link between self-efficacy and academic performance [21, 39, 53]. The findings are not entirely consistent. However given the precise meaning of the concept of self-efficacy, the relationship between perceived self-efficacy and academic performance is more likely to be apparent when the inventory used in the investigation is specific to the content area of interest [21, 39].

2 Measuring Perceived Statistics Self-Efficacy

Many researchers have attempted to measure students’ perceived self-efficacy in relation to statistics coursework. As early as 1989, Benson (1989) conducted a study investigating factors related to statistics test anxiety [10]. He found that that self-efficacy was not related to graduate or undergraduate students' performance in their required statistics coursework. Benson's research incorporated only three questions about self-efficacy which were not specific to the performance tasks inherent in statistics coursework. Bandalos, Yates and Thorndike-Christ (1995) also concluded that there is no relation between students' performance in their statistics coursework and their perceived self-efficacy [6]. Bandalos, Yates and Thorndike-Christ (1995) developed seven items intended to measure statistics self-efficacy, asking students to respond to questions such as how they felt about their ability to "construct a graph". Bandalos, Yates and Thorndike-Christ (1995) posit that the items developed may have failed to demonstrate a specific connection to statistics as the "self-efficacy" items were found to be stronger indicators of math self-concept, another construct they included in their research. Awang-Hashim et al. (2002) incorporated a general trait assessment in an effort to study the relationship of self-efficacy, effort, and worry to student performance in statistics. Students were asked to respond with their feelings regard to eight questions such as "I'm confident I can learn the most complex material presented by the instructor in this course" [4]. Again, no significant relation was found between self-efficacy and the students' performance in this investigation. However each of these studies seems to fail to appreciate the situation-specific nature of the construct of self-efficacy and operationalize it instead as a general dispositional construct. As indicated by a number of researchers [21,39], if the goal is to measure self-efficacy specifically related to statistics, this will be more precisely accomplished with an inventory asking students to respond with their beliefs about their ability to perform statistical tasks rather than an inventory constructed to assess mathematics-related self-efficacy. According to Pajares (1996), "studies that report a lack of relation between self-efficacy and performance often suffer from problems in …specificity". More recently Lane et al. (2004) developed the self-efficacy towards statistics questionnaire (STSQ) [29]. This inventory includes 44 items asking students about their beliefs about their ability to perform statistics-related tasks as well as items asking students about their beliefs about their general behavior such as "how confident are you in your ability to organize your time?" or "how confident are you about your ability to motivate yourself". Their results suggested that students' scores on the STSQ were only moderately correlated with the students' final class performance. However in light of the "task-specific" nature of self-efficacy and some of the more general items included in the STSQ, it is questionable if the STSQ genuinely measures the construct of self-efficacy specifically related to statistics. Presently this instrument has been validated only by the developers.

3 Current Statistics Self-Efficacy

Finney and Schraw (2003) developed two forms of their inventory intended to measure self-efficacy specifically related to statistics: the current statistics self-efficacy (CSSE) inventory and the self-efficacy to learn statistics (SELS) [17]. These inventories followed the structure of earlier math self-efficacy
inventories (e.g., MSES-R, Kranzler & Pajares, 1997) but with the addition of specific terminology related to particular concepts within the statistics domain [27]. The SELS was developed for students at the beginning of their statistics course experience in an effort to measure their perceived ability to learn statistics. The CSSE was developed to assess the students’ confidence with respect to their current ability to solve specific tasks related to statistics. The 14 items, on both of these inventories, are questions specific only to performing tasks in statistics. Both inventories have been validated by Finney and Schraw (2003) and were found to have acceptable levels of internal consistency. Measured with Cronbach’s α, internal consistency was moderate for the SELS and high for the CSSE. Finney and Schraw report that scores on the CSSE were significantly related to students’ final course grades. Abd-El-Fattah (2005) used the CSSE in research examining current statistics self-efficacy (Finney & Schraw, 2003) in a computer-based statistics course and found that student’s scores on the CSSE were significantly correlated to the students’ final performance in the course [1,17]. Abd-El-Fattah's study is the only available research found in the current literature using the Finney and Schraw inventory. Prior to the current investigation, studies incorporating confirmatory factor analyses to investigate the construct validity of this inventory do not currently exist in the research literature regarding statistics education.

4 Mathematics Education

Research in statistics education has routinely explored the possible connection between mathematics and the psychological constructs related to the process of learning statistics. Gal, Ginsburg, and Schau (1997) found that many students believe that statistics is simply a continuation of mathematics. The students they interviewed possessed beliefs that statistics coursework required difficult mathematical knowledge, good computational ability and abstract thinking [20]. Although students’ perceptions do impact how they approach the subject area of statistics, mathematics and statistics are not one in the same. Statistics is not a subcategory of mathematics but rather statistics is a science much like physics is a science. Although statistics depends heavily on mathematical computations, statistics is “the science of inference” with different modes of thinking and distinctive concepts from mathematics (Moore, 1997) [34]. This is further established by research demonstrating that statistics and mathematics differ in the cognitive processes necessary for each area. According Cruise et al. (1985), statistics involves different mental processes and is more than the manipulation of mathematical symbols [15]. Fox et al. (1979) argue that statistics involves reasoning skills that more closely resemble verbal reasoning skills than mathematical reasoning skills and according to Fox, et al., solving statistical problems involves more logical skills than mathematical skills [19]. Birenbaum and Eylath (1994) demonstrated that the cognitive processing associated to statistics differs from that associated with mathematics [13]. They found that student's negative feelings in statistics were significantly related to students' inductive reasoning ability as well as their numerical processing ability. However student's negative feelings about mathematics were only found to be significantly related to numerical processing ability. Birenbaum and Eylath's findings are consistent to the positions of Cruise et al. (1985) and Fox et al (1979) presented earlier. Thus in spite of their strong relations, statistics and mathematics are not the same and subsequently student's feeling about statistics and mathematics also differs. Students with high levels of negative affect about mathematics tend to have high levels of negative affect about statistics but the converse in not necessarily true [36]. There is evidence that student cognition in the area of mathematics may be related to the student's statistics related self-efficacy [17]. Finney and Schraw (2003) demonstrated a strong relation between students' mathematics self-efficacy and statistics related self-efficacy to three different groups of students taking a statistics class. However students' CSSE scores were still more strongly related to final course grade than they were to relate to students' mathematics self-efficacy scores. This is not unexpected. As stated earlier self-efficacy is a context-specific construction and failure to include "specificity" will result in incorrect conclusions [39]. Measured
specifically, self-efficacy should predict statistics performance better than mathematics self-efficacy. Therefore mathematics self-efficacy and statistics self-efficacy while related are not the same construct.

5 Current Study

The goal of the current investigation was to test the one-dimensional construct of statistics related self-efficacy as measured with the CSSE [46]. CFA is the appropriate analysis when the researcher has some knowledge of the underlying latent variable structure, either by method of prior empirical research or by method of theoretical knowledge. Therefore CFA was used to demonstrate the construct validity of the CSSE by testing the degree of correspondence between the one-dimensional theoretical construct and the observed data. The CSSE was used to measure participants’ statistics-related self-efficacy [39]. This instrument was chosen in light of the fact that students in the present investigation are graduate-level students and therefore, had some prior experience with statistics, either in research methods classes, prior statistics classes or during independent research experiences. Many of the questions in the CSSE reference statistical terms that would only be familiar to a student with some prior experience of statistics education. This will be the first study conducting a confirmatory factor analysis on the CSSE. In light of the research suggesting a relation among mathematics cognition, affect and performance to statistics cognition, affect, and performance, the present study also uses SEM to examine the theorized causal link between students’ prior mathematics experiences and statistics–related self-efficacy. This prior mathematics experiences construct includes self-reported measures of students’ high school mathematics experience; student’s reported mathematics self-concept and prior experience with statistics. This will be the first study that demonstrates a causal link between student’s perceptions of their earlier mathematics experiences with their current statistics-related self-efficacy. Finally, post SEM analyses are used to examine the effect of demographic variables such as gender, age, ethnicity, department, degree level, hours completed, expected course grade, number of college-level math classes, and current GPA, on the latent factor statistics-related self-efficacy, after extracting latent variable values after associated error has been isolated out of the variable [25]. These variables were chosen as they are all commonly found variables in existing statistics education research examining other psychological constructs [9, 22, 36, 50, 55, 58]. Since research in the area of self-efficacy, specifically related to statistics coursework is limited, this study is the first that will examine these ancillary variables in relation to this psychological construct. Recently the use of structural equation modeling has advance to providing researchers with a method of extracting latent variable values after associated error has been isolated out of the variable. Since the error variance has been factored out, the resulting factors scores are essentially error free. Because latent variable scores are error-free measures, these values offer increased precision in detecting true relations between theoretical concepts measured by observed variables. The exact procedures for calculating such scores are outlined by Jöreskog (2000). Latent variable scores are easily computed in LISREL® 8.8 [28]. Currently there are no studies available in the existing literature that demonstrates the use of latent variable scores in post structural analyses, as suggested can be conducted in Jöreskog (2000). The use of these procedures adds a great deal of flexibility to the potential analyses that can be performed. These analyses provide the mechanism for examining if these ancillary variables (gender, age, ethnicity, department, degree level, hours completed, expected course grade, number of college-level math classes, and current GPA) effect the latent construct (statistics related self-efficacy) with a sample size that prohibits the use of multi-group SEM analyses.
6 Method

6.1 Participants
The participants included 238 graduate level students enrolled in statistics courses offered in the departments of biological science, education, geography, and psychology at a large Mid-Western University. This sample of students included both master and doctoral level students. These three credit hour graduate-level courses are intended to cover univariate and bivariate techniques. Students participated during the fall, spring, and summer semesters. These courses were taught by different instructors. Seventeen students chose not participate in the survey, resulting in an overall response rate of 93.3%. In order to participate, students were required to complete and sign the IRB required consent form. The ages of the participants ranged from approximately 23 to 61, with a majority of the participants being female, \( n = 176 \) (73.9 %) and Caucasian-American, \( n = 165 \) (69.3 %).

6.2 Instrumentation
The current statistics self-efficacy (CSSE) inventory is an instrument developed by Finney and Schraw (2003) to assess as a one-dimensional construct. With this instrument, respondents are asked to rate their current belief in their ability to complete 14 specific tasks related to statistics using a 1 (no confidence at all) to 5 (complete confidence) response scale. Questions included such statements as "identify a scale of measurement for a variable", "interpret the results of a statistical procedure in terms of the research question” and "identify the factors that influence power”. The CSSE is generally completed by students in less than five minutes. A complete list of the CSSE questions is provided in Appendix A. Additionally students were also asked to provide some basic demographic information regarding themselves and their prior mathematics experiences. The prior mathematics experiences questions include three questions asking students to respond to the following statements, using a five-point scale: 1) "how well did you do in your high school mathematics courses?"; 2) "how good at mathematics are you?"; and 3) "how much experience with mathematics have you had?". These questions have been used by a number of researchers to understand the prior experiences of their students [46, 49].

6.3 Procedure
The present investigation involves survey research by means of convenience sampling, a form of non-probability sampling [57]. Students were approached during the fourth week of class once permission was given by the section instructor. Students were informed that data was being collected on the instrument in an effort to investigate how they feel about the statistics/quantitative methods coursework that they were required to successfully complete. Participants were informed that they would be required to sign a consent form in order to participate and that they would be paid five dollars cash for their participation once they completed the survey instruments. On the consent form, students were informed that study findings would be made available to all interested participants once data collection and analysis has been completed. Students were encouraged to fill out the survey as completely as possible, and to ask questions if they did not understand something on the instruments. After preliminary information regarding the project was explained, the survey was distributed to the students in file folders with two copies of the required consent forms. Upon opening the file folder, students were directed to sign one consent form and to keep the duplicate consent form for their records. Pre-assigned numbers were placed on the instruments in an effort to protect student's identities and so that demographic information could be matched with the completed CSSE survey. Survey data was electronically scanned (three times) into the statistics package for the social science (SPSS, 2003) in order to reduce error involved with hand entry.
7 Data Preparation

Once data was collected, a number of procedures were used to prepare the data for subsequent analyses. First the data was examined for missing values. Less than one percent of the total item-responses were not completed in the entire data set. Since there was no pattern to the missing responses, multiple imputation procedures, generated through the linear structural relations program (LISREL® 8.8, 2006), were used to complete the sixteen missing responses. Multiple imputations are one of many methods available for dealing with missing data (Fox-Wasylyshyn et al., 2002) [19]. Multiple imputation was implemented in the present study because it is considered by many researchers to be the superior approach to dealing with missing data [2,16,26,44,45] and unlike other methods, multiple imputation has been found to be robust to model violations [2,26]. The procedure is readily available in the LISREL program. Multiple imputations are accomplished through several stages of data analyses in which data from complete cases is used to predict the value of the missing item.

8 Threshold Analyses

Bollen (1989) argued that one of the most frequent violations in the use of SEM is the failure of researchers to make necessary adjustments to non-continuous data [12]. However more than twenty-five years later many studies, including many of the studies found in research on statistics education that use SEM fail to make these appropriate data adjustments [5, 23, 35, 49]. For the present investigation, all survey items were Likert-type scale items, and therefore are considered to be ordinal in scale. These responses were transformed into continuous data through threshold analyses. In contrast, the current standard in research is to treat Likert-scale, ordinal variables as if the values have metric properties similar to continuous variables [12, 18, 24, 31]. However this approach to data analyses is problematic for the reason that Likert-scale response choices do not have a unit of measurement. Also data analyzed through structural equation modeling techniques are generally estimated by method of maximum likelihood (ML) techniques which assume that data is continuous and normally distributed. Ordinal variables, which are categorical observations, are neither continuous nor normally distributed. The use of non-normal data in SEM can introduce errors associated with heteroscedasticity that occur to the use of ordinal measures [31]. Therefore it is not appropriate to use ordinal-type data in structural models; and instead it is necessary to produce thresholds in order to establish meaningful correlation, covariance, and mean matrices from the ordinal data [24]. Originating with McKelvey and Zavoina (1975), threshold analyses are founded on the postulation that for each ordinal variable Y, there exists an underlying continuous variable Y* in which \(-\infty < Y* < \infty\) [44]. This theorized continuous variable represents the affect underlying the ordered responses to Y in such a way that the indices no longer have the "...problem of equal weights for all possible answers" [40]. For the present study, this underlying distribution was established through alternative parameterization since this approach is considered the gold standard when all items have an equal number of response choices and the researcher is interested in all threshold values that are positive values. A complete discussion of threshold analyses can be found in Jöreskog (2000) and Perakis, Maravelakis, Perakis, & Xekalaki (2005).

9 Results

9.1 Reliability

A Cronbach's alpha was calculated for the data collected in an effort to analyze the internal consistency of items in the inventory [14]. Cronbach's alpha is the appropriate reliability assessment as the construct of was being tested as a one-dimensional construct. This analysis is be conducted on the ordinal responses, revealed acceptably high levels of reliability with \( \alpha = .917 \) on the 14 items of the CSSE inventory [52].
9.2 Demographic Information

Descriptive analyses were conducted on the demographic information provided by participants. The analyses revealed that there was little variance in students’ grades and expected grades at the graduate level: reported grade point averages (GPA) were $M = 3.83, SD = 0.25$, with values ranging from 3.0 to 4.0; expected final grades for their present statistics/quantitative methods classes were $M = 3.78, SD = 0.38$, with values ranging from A to a B-. This indicates that 96% of students expected to receive a final grade of 3.0 (B) or higher. Due to this lack of variability, no analyses in the present study examined student achievement in relation to students CSSE scores. Demographic information for the participants is presented in Table 1.

Table 1: Demographic characteristics of sample participants ($N=238$)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>$n$</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at time of survey (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>135</td>
<td>56.7</td>
</tr>
<tr>
<td>30-39</td>
<td>62</td>
<td>26.1</td>
</tr>
<tr>
<td>40-49</td>
<td>27</td>
<td>11.3</td>
</tr>
<tr>
<td>50-59</td>
<td>13</td>
<td>5.5</td>
</tr>
<tr>
<td>60-69</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian American</td>
<td>165</td>
<td>69.3</td>
</tr>
<tr>
<td>African American</td>
<td>20</td>
<td>8.4</td>
</tr>
<tr>
<td>Hispanic American</td>
<td>4</td>
<td>1.7</td>
</tr>
<tr>
<td>Asian American</td>
<td>7</td>
<td>2.9</td>
</tr>
<tr>
<td>Other American</td>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>Foreign Student</td>
<td>36</td>
<td>15.1</td>
</tr>
<tr>
<td>Department</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>114</td>
<td>47.9</td>
</tr>
<tr>
<td>Psychology</td>
<td>55</td>
<td>23.1</td>
</tr>
<tr>
<td>Geography</td>
<td>19</td>
<td>8.0</td>
</tr>
<tr>
<td>Business</td>
<td>7</td>
<td>2.9</td>
</tr>
<tr>
<td>Biological Science</td>
<td>37</td>
<td>15.5</td>
</tr>
<tr>
<td>Communication</td>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>Degree Currently Seeking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masters</td>
<td>87</td>
<td>36.6</td>
</tr>
<tr>
<td>Doctorate</td>
<td>142</td>
<td>59.7</td>
</tr>
<tr>
<td>Specialist</td>
<td>6</td>
<td>2.5</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>1.2</td>
</tr>
<tr>
<td>Hours Earned Toward Present Degree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-20</td>
<td>159</td>
<td>66.8</td>
</tr>
<tr>
<td>21-40</td>
<td>45</td>
<td>18.9</td>
</tr>
<tr>
<td>41-60</td>
<td>12</td>
<td>5.0</td>
</tr>
<tr>
<td>61-80</td>
<td>12</td>
<td>5.0</td>
</tr>
<tr>
<td>81-100</td>
<td>10</td>
<td>4.2</td>
</tr>
</tbody>
</table>
Students were also asked to respond to additional questions regarding their prior mathematics experiences. The distributions of students' responses are presented in Table 2.

<table>
<thead>
<tr>
<th>Question and scale</th>
<th>% of responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>How good are you at mathematics? (very poor to very good)</td>
<td>3.8% 8.4% 18.9% 36.1% 32.8%</td>
</tr>
<tr>
<td>How well did you do in your high school math course? (very poorly to very well)</td>
<td>4.2% 6.3% 18.1% 39.5% 31.9%</td>
</tr>
<tr>
<td>How much experience have you had with mathematics? (none to a great deal)</td>
<td>0% 22.2% 30.3% 34.5% 13.0%</td>
</tr>
</tbody>
</table>

As can be seen by the distribution of the responses in Table 2, with the exception of student responses to the question "how much experience have you had with mathematics", responses were all positively skewed. The failure of Likert-scale responses to be normally distributed and the fact that such ordered data have no common origin requires that these responses be transformed (via threshold analyses conducted) prior to empirical analysis in structural equation modeling. Finally students also were asked to indicate how many college-level mathematics classes they had taken. Students reported taking an average of four college-level mathematics classes (\(M = 4.07, SD = 4.133\)) with a range of zero to thirty-eight classes reported.

### 9.3 Factor Loadings and Reliability of the Statistics-Related Self-Efficacy Construct

Confirmatory factor analysis (CFA) was conducted on the 14 item CSSE data in an effort to establish the construct validity of the instrument with the current population [59]. For these analyses, item EF1 was fixed to a value of one for the purposes of model convergence. As revealed in these analyses, self-efficacy is a one-dimensional factor which converges on the 14 items from the CSSE inventory. The model demonstrated good fit (\(\chi^2 = 598.64, p<.001, CFI = 0.948, NNFI = 0.938, RMSEA = 0.127\)) according to the guidelines proposed by Schermelleh-Engel, Moosbrugger and Müller (2003). Results reveal that all factor loadings between the first-order factor of self-efficacy and its associated items were all significant at \(p < .05\). Therefore since the theoretical model of has been verified with the current sample of data, it is reasonable to conduct additional analyses with this verified model. Reported parameter estimates are raw coefficients. The standardized coefficients are not reported here because they are square roots of item's reliabilities included in Table 3.
Table 3: First-order factor loadings and squared-multiple correlations (R²) for statistics-related self-efficacy (A)

<table>
<thead>
<tr>
<th>Item</th>
<th>Parameter estimate</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF1</td>
<td>0.597*</td>
<td>0.357*</td>
</tr>
<tr>
<td>EF2</td>
<td>0.675*</td>
<td>0.455*</td>
</tr>
<tr>
<td>EF3</td>
<td>0.780*</td>
<td>0.609*</td>
</tr>
<tr>
<td>EF4</td>
<td>0.769*</td>
<td>0.591*</td>
</tr>
<tr>
<td>EF5</td>
<td>0.763*</td>
<td>0.582*</td>
</tr>
<tr>
<td>EF6</td>
<td>0.617*</td>
<td>0.381*</td>
</tr>
<tr>
<td>EF7</td>
<td>0.746*</td>
<td>0.556*</td>
</tr>
<tr>
<td>EF8</td>
<td>0.692*</td>
<td>0.480*</td>
</tr>
<tr>
<td>EF9</td>
<td>0.782*</td>
<td>0.612*</td>
</tr>
<tr>
<td>EF10</td>
<td>0.752*</td>
<td>0.566*</td>
</tr>
<tr>
<td>EF11</td>
<td>0.795*</td>
<td>0.632*</td>
</tr>
<tr>
<td>EF12</td>
<td>0.659*</td>
<td>0.435*</td>
</tr>
<tr>
<td>EF13</td>
<td>0.669*</td>
<td>0.447*</td>
</tr>
<tr>
<td>EF14</td>
<td>0.656*</td>
<td>0.430*</td>
</tr>
</tbody>
</table>

Note: *p < .05 and bolded item was fixed to 1.0.

These values indicate the percent of the variance in the observed variables attributed to the first-order factor. These values are given as the squared multiple correlation coefficient for each item and are a variance ratio (true score to observed score) for a specific item or latent construct and indicate the degree to which each item adequately measures its respective underlying construct [12].

9.4 Congeneric Model of Prior Mathematics Experiences

As the foregoing review of the literature suggests, many researchers have proposed that students’ prior mathematics experiences can be predictive of student’s statistics-related self-efficacy. In an effort to examine the link between prior mathematics experiences and the14-item model of CSSE, it was necessary to develop the factor of prior mathematics experiences. The prior mathematics experiences construct was composed of the student responses to the three questions: 1) "how well did you do in your high school mathematics courses?"; 2) "how good at mathematics are you?"; and 3) "how much experience with mathematics have you had?" The single-factor model with these variables revealed that the three variables were significant measures in the underlying factor, prior mathematics experiences. As revealed in Table 4, the three items adequately measured the underlying construct and prior mathematics experiences.

Table 4: Parameter estimates and squared-multiple correlation of prior mathematics experiences causal factor

<table>
<thead>
<tr>
<th>Item</th>
<th>Parameter Estimate</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>How well did you do in your high school mathematics courses?</td>
<td>0.758**</td>
<td>0.215*</td>
</tr>
<tr>
<td>How good at mathematics are you?</td>
<td>0.814**</td>
<td>0.149*</td>
</tr>
<tr>
<td>How much experience with mathematics have you had?</td>
<td>0.587**</td>
<td>0.266*</td>
</tr>
</tbody>
</table>

Note: *p < .05; **p < .01; Reliability is squared-multiple correlation.

9.5 Causal Model of Prior Mathematics Experiences on Statistics-Related Self-Efficacy

It is hypothesized that prior mathematics experiences a direct predictor of each of statistics-related self-efficacy. The γ-coefficient, linking statistics-related self-efficacy to prior mathematics experiences was positive and significant; the p-value was less than α= 0.05. Prior mathematics experiences on self-efficacy demonstrated a reasonable fit (χ² =491.32, p <.001, CFI = 0.963, NNFI= 0.961, RMSEA = 0.116). In the
model in which prior mathematics experiences factor was introduced as a predictor of the primary factor, statistics-related self-efficacy revealed that prior mathematics experiences accounted for a significant 56.2% (squared multiple correlation, $R^2 = .562$) of the variance in statistics-related self-efficacy.

9.6 Results of Latent Variable Analysis

Analysis of variance (ANOVA) was conducted on the exogenous variables of gender, age, ethnicity, department, degree seeking, hours, expected course grade, current GPA, and number of college-level mathematics classes in order to analyze the effect of each of these variables on the students CSSE score. A summary of these ANOVA is presented in Table 5.

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>df</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>236</td>
<td>0.039</td>
</tr>
<tr>
<td>Age</td>
<td>4</td>
<td>233</td>
<td>1.810</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>5</td>
<td>232</td>
<td>1.550</td>
</tr>
<tr>
<td>Department</td>
<td>5</td>
<td>232</td>
<td>6.18*</td>
</tr>
<tr>
<td>Degree</td>
<td>4</td>
<td>233</td>
<td>0.701</td>
</tr>
<tr>
<td>Hours</td>
<td>4</td>
<td>233</td>
<td>0.320</td>
</tr>
<tr>
<td>Expected Grade</td>
<td>5</td>
<td>237</td>
<td>6.14*</td>
</tr>
<tr>
<td>GPA</td>
<td>1</td>
<td>236</td>
<td>0.335</td>
</tr>
<tr>
<td>Number of Classes</td>
<td>7</td>
<td>230</td>
<td>3.70*</td>
</tr>
</tbody>
</table>

Note: *$p<0.05$

As indicated in Table 5, the exogenous variables department, expected grade, and number of college-level mathematics classes revealed significant differences in the students' CSSE scores. There were no significant differences in students' CSSE scores for the remaining exogenous variables.

9.6.1 Department

The department item asked students to indicate in what department in the university they were associated. The department variable included six departments: biological sciences ($n = 37$), business ($n = 7$), communication ($n = 6$), education ($n = 114$), geography ($n = 19$) and psychology ($n = 55$). The one-way ANOVA analyses indicated that there was a significant difference in reported statistic-related self-efficacy with regards to department, $F (5, 232) = 6.18, p < .001$. Scheffe post hoc analyses indicated that the mean of self-efficacy scores for students from biological sciences was 0.794 lower than that reported by psychology students and 0.595 lower than that reported by education students.

9.6.2 Expected Course Grade

Students were asked to indicate what grade they expected to get in their present statistics/quantitative methods course as well as their present GPA. Students expected grades for the class resulted in the following categories: A ($n = 140$), A- ($n = 47$), B+ ($n = 17$), B ($n = 23$), B- ($n = 7$), and C's and below ($n = 4$). The one-way ANOVA analyses indicated that there was a significant difference in statistics-related self-efficacy regards to expected course grade, $F (5, 237) = 6.14, p < .001$. Scheffe post hoc analyses revealed that significantly higher mean scores in reported self-efficacy for students expecting an A relative to students expecting a B ($ΔM = .545$).

9.6.3 Number of College-level Mathematics classes

Students were asked to indicate how many college-level mathematics classes that they had taken. This variable was categorized accordingly: 0-1 ($n = 38$); 2 courses ($n = 44$), 3 courses ($n = 55$), 4 courses ($n =
34), 5 courses (n = 27), 6 courses (n = 13), 7 to 9 courses (n = 14) and 10 or more courses (n = 13). The one-way ANOVA analyses indicated that there was a significant difference in reported level of statistics-related self-efficacy with regards to number of college-level mathematics classes, \( F(7, 230) = 3.70, p < .001 \). Specifically students reporting more college-level mathematics classes had significantly higher self-efficacy scores. Scheffe post hoc analyses revealed significantly lower mean scores in reported self-efficacy for students who had taken two courses relative to those who had taken 5 courses (\( \Delta M = .654 \)).

10 Discussion

Taken statistics courses with some pre-conceived ideas that often create apprehension about their ability to be successful. Bandura (1997) suggests that students with higher levels of perceived self-efficacy towards a particular subject area will be more likely to engage the information, participate in class, persist though difficulties and achieve success at a higher level, relative to students who do doubt their capabilities with the particular coursework. The current study examines the factors that influence student's statistics-related self-efficacy that will be present while completing their required statistics coursework.

10.1 Confirmatory Factor Analysis

The data collected from the present sample of participants' demonstrated a high level of reliability. Additionally construct validity of the CSSE inventory was established through strict confirmatory factor analyses, revealing the one-dimensional construct proposed by Finney and Schraw (2003) while demonstrating good model fit. Since confirmatory factor analyses indicated that the CSSE had been appropriately measured, it was reasonable to use this data to address the research questions guiding the present investigation. This study is the first study establishing the construct validity of the CSSE.

10.2 Causal Model of Prior Mathematics Experiences and Statistics-Related Self-Efficacy

The causal model, exploring the link between prior mathematics experiences and statistics-related self-efficacy as measured by the CSSE that demonstrated good fit. Consistent with a number of studies in the research on statistics education, these finding suggest that students' prior mathematics experiences are a factor in the students' statistics-related self-efficacy scores. The model suggested that an overwhelming 56.2% of the variance in statistics-related self-efficacy can be attributed to prior mathematics experiences. This is the first study that has demonstrated a causal link of student's perceptions of their prior mathematics experiences and student's current statistics related self-efficacy. These findings suggest that student's statistics-related self-efficacy is strongly influenced by their perceptions of their prior mathematics experiences.

10.3 Latent Variable Analysis: Significant Effect of Department

One-way ANOVA analyses revealed a significant effect of students' department on self-efficacy scores. Specifically scores were significantly lower for students from the department of biological sciences relative to students in the departments of psychology and education. Why such a pattern of results might exist is not intuitively obvious. Examination of cross-tabulation analyses (ethnicity by department) indicated a significantly higher percentage of students reporting to be "international" in the biological sciences department (n = 13, 35.1%) relative to psychology (n = 9, 16.5%) and education (n = 9, 7.9%) departments. It is possible that these results may be due to cultural differences that exist in levels of self-efficacy. Cultural differences in self-efficacy have been suggested as a possibility by other researchers in the past; however, conclusive research in this area is presently not available [30, 56]. Another possibility is that these results reflect that fact that 76% of the biological sciences students were first semester graduate students. Students in education and psychology included both master level and doctoral level students. The
presents of such a high percentage of international students in a new master program may shed some light as to these significant differences.

10.4 Latent Variable Analysis: Number of Prior College-Level Mathematics Classes
Consistent with the suggestions of prior research, the number of college-level mathematics courses taken was found to be significantly related to students’ level of statistics-related self-efficacy. These results indicate that students having more college-level mathematics classes had significantly higher statistics-related self-efficacy. Unlike earlier studies, this current investigation was able to demonstrate this conclusion with the use of essentially error-free latent-variable scores. This superior approach to data analysis significantly reduces the potential of conclusions rampant with statistical error (Type II) in these analyses.

10.5 Latent Variable Analysis: Expected Grade in Present Statistics Class
Consistent with several prior research [3,54], students’ expected grade was found to be significantly related to students’ level of statistics-related self-efficacy. As anticipated students expecting A’s had higher levels of statistics related self-efficacy, relative to students who were expecting B’s. These findings are consistent to "self-efficacy" in that those who would believe that they could be most successful in their required statistics coursework are more likely to report an expected final grade of "A" relative to those who might not have as high of a level of statistics-related self-efficacy. Again it is worth noting that the sample of participants was limited to graduate-level students; who are under a greater level of pressure (real or imagined) to get A’s.

10.6 Latent Variable Analysis: Gender, Age, Ethnicity, GPA, Degree Seeking and Hours Completed
The analyses of the factor scores revealed no significant differences in and the remaining variables of gender, age, ethnicity, overall GPA, degree seeking and credit hours completed. Although these exogenous variables are those commonly examined in the research in statistics education, this present investigation is the first to present research examining the effect of these variables on self-efficacy specifically related to statistics education. The impact of gender has been one of the most widely researched variables in the area of statistics education; however, no studies have looked at statistics-related self-efficacy as a dependent measure. Findings have generally been mixed as to whether gender is a factor in experiences and achievement associated with statistics coursework [16,23,33,43,59]. It is possible that the failure to find significant differences in the present investigation is due to the fact that the sample of students used was comprised of graduate students. Males and females at the graduate level may represent unique subsets of their respective genders that share more academic similarities rather than differences often associated with the opposite sex. Age is another common factor in statistics education research; however no studies have examined age as a factor in relation to statistics-related self-efficacy. Again results in the statistics education research are somewhat mixed with older students generally fairing poorer on performance measures relative to younger students [43,55]. In the present study, the greatest range in ages was found only in the department of education; students in the other departments were generally in their twenties. Again the current study found no significant relationship between level of statistics-related self-efficacy and student’s reported age. There are very few studies that have investigated the role of ethnicity in statistics education. Some have found no relation between ethnicity and performance [36,50,59]. However Bell’s (1998) study of statistics-related Anxiety found that international students in spite of stronger mathematics backgrounds, reported higher levels of statistics-related anxiety. This might go along with the current findings. Although no significant differences were found for the ethnicity variable, the department with the greatest international student representation (biological sciences) did report significantly lower statistics-related self-efficacy. The association between self-efficacy and ethnicity needs further examination in subsequent research. Currently no studies have found a significant effect of overall GPA in
the research in statistics education [22,36,42] consistent to the current investigation. Additionally the present investigation found no significant effects of the degrees students were presently working towards (e.g., master, doctoral, et cetera) on their level of statistics-related self-efficacy. In one prior study, Benson and Bandalos (1989) found that graduate students had significantly higher statistics anxiety than undergraduates. However no further analyses were conducted which indicated if there were significant differences among graduate students seeking different degrees. Again subsequent research should examine the relation among statistics-related self-efficacy, statistics-related anxiety and how these impact students at different stages in their academic careers.

11 Conclusion

The current investigation examined the theoretical construct of statistics-related self-efficacy as measured with the CSSE (Finney & Schraw, 2003) and the relation of this inventory to variables generally analyzed in statistics education research. The findings suggest that prior mathematics experiences do indeed impact the graduate student's belief about their abilities to perform tasks associated with their required statistics coursework. Additionally the findings of the current investigation suggest that students with more college-level mathematics classes completed also reported higher levels of statistics-related self-efficacy. These results are important as studies reveal that up to 80% of the statistics students surveyed continue to experience problems with and have bad attitudes towards statistics [37,38]. The predictive power of self-efficacy illustrates that it could serve a useful function in identifying students at risk of failure in the early stages of a statistics coursework. Potentially these results suggest that students with weaker backgrounds in mathematics could benefit from some remedial work in the area of mathematics. These results might also suggest that there may be some benefit to computer-assisted instruction for these students since feelings about mathematics are often transferred to statistics coursework [41]. The use of computer-assisted instruction may be one approach to minimizing the degree to which student's transfer feelings about mathematics to their required statistics coursework. Additional research, examining statistics-related self-efficacy, in computer/software-assisted statistics education could potentially shed light on whether the carryover of student's feelings about mathematics to their statistics coursework can be alleviated through the incorporation of technology into the classroom [42]. Lastly the current study demonstrates the use of latent variable analyses, post SEM procedures. While the current study tested the construct validity of the one-dimensional factor of statistics-related self-efficacy proposed by Finney & Schrow (2003) the n = 238 sample size made it impossible to test multi-group models that would reveal the effect of variables such as gender, age, ethnicity, department, degree level, hours completed, expected course grade, number of college-level math classes, current GPA that have been examined in other area of statistics education. The impact of these variables has not been examined specifically in reference to statistics-related self-efficacy. The somewhat neglected procedure presented originally by Jöreskog (2000) provided the mechanism by which the effect of a number of previously untested variables on student's statistics-related self-efficacy could be examined.

Appendix

CSSE Self-Efficacy Items

1) Interpret the probability value from a statistical procedure.

2) Identify if a distribution is skewed when given the values of three measures of central tendency.

3) Select the correct statistical procedure to be used to answer a research question.
4) Explain what the value of the standard deviation means in terms of the variable being measured.

5) Distinguish between a Type 1 error and a Type 2 error in hypothesis testing.

6) Explain what the numeric value of the standard error is measuring.

7) Distinguish between the information given by the three measures of central tendency.

8) Distinguish between a population parameter and a sample statistic.

9) Explain the difference between a sampling distribution and a population distribution.

10) Identify a scale of measurement for a variable.

11) Interpret the results of a statistical procedure in terms of the research question.

12) Identify the factors that influence power.

13) Distinguish between the objectives of descriptive versus inferential statistical procedures.

14) Identify when the mean, median, and mode should be used as a measure of central tendency.

References

http://dx.doi.org/10.1177/0049124100028003003


http://dx.doi.org/10.1207/S15326977EA0804_3


http://dx.doi.org/10.1037/0022-0663.87.4.611


http://dx.doi.org/10.1080/0013188940360110


http://dx.doi.org/10.1007/BF02310555


http://dx.doi.org/10.1177/1094428103255532


http://dx.doi.org/10.1016/S0361-476X(02)00015-2


http://dx.doi.org/10.1037/1082-989X.9.4.466


http://dx.doi.org/10.1207/s15326985ep3902_3

http://dx.doi.org/10.1016/S0191-8869(01)00148-9


http://dx.doi.org/10.1089/cpb.2006.9.441


http://dx.doi.org/10.1111/j.2044-8279.1991.tb00989.x

http://dx.doi.org/10.1080/00223891.1990.9674054