www.ijres.net

Re-Validation and Exploration of Modified Versions of the Statistics Anxiety Scale Developed for College Students in the United States

Keston G. Lindsay
University of Colorado Colorado Springs, United States

Jessica B. Kirby
University of Colorado Colorado Springs, United States

Brynn C. Adamson
University of Colorado Colorado Springs, United States

To cite this article:

Lindsay, K.G., Kirby, J.B., & Adamson, B.C. (2024). Re-validation and exploration of modified versions of the Statistics Anxiety Scale developed for college students in the United States. *International Journal of Research in Education and Science (IJRES), 10*(2), 426-445. https://doi.org/10.46328/ijres.3400

The International Journal of Research in Education and Science (IJRES) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

IJRES International Journal of Research in Education and Science
2024, Vol. 10, No. 2, 426-445 https://doi.org/10.46328/ijres.3400

Re-Validation and Exploration of Modified Versions of the Statistics Anxiety Scale Developed for College Students in the United States

Keston G. Lindsay, Jessica B. Kirby, Brynn C. Adamson

Article Info

Article History

Received:
20 December 2023
Accepted:
30 April 2024

Keywords

Statistics anxiety
Statistics anxiety scale
Statistics education

Abstract

This study aimed to validate a modified Statistics Anxiety Scale for students in the United States taking university courses. Modifications were made by changing the wording of several items to be consistent with American English, and to accommodate students taking statistics courses in various formats. Items were added to investigate anxiety toward the use of statistical packages, and peer mentoring. Data from 352 participants and exploratory factor analyses were used to analyze the original 24-item SAS (SAS-O) and a version of the SAS with six additional items (SAS-M). The three-factor structure for SAS-O was consistent with the original validation study, explaining about 64 % of the items' variance. The factor structure for SAS-M contained an additional two factors, that explained a total of 68 % of the items' variance. The factors were internally consistent, correlated with one another, and negatively correlated with Wise's Attitude Toward Statistics scale. Male students generally had lower application anxiety and examination anxiety than female students, and lower asking for help anxiety than non-traditional students.

Introduction

There is much literature demonstrating the importance of experimental design and statistical application in STEM fields such as engineering (Zhan, Fink, & Fan, 2010), biology (Pollard et al., 2019), chemistry (Hartland, 2020), and physics (Pfaff et al., 2013), and in applied fields such as kinesiology (Beck, 2013; Hopkins et al., 2008), nursing (Epstein et al., 2011; Hanoch & Pachur, 2004), and public health (Hayat et al., 2017). According to the Bureau of Labor Statistics (BLS, 2022), statistics has been cited as an important component for problem-solving and research in the engineering and scientific fields, and the job outlook for statisticians is projected to increase up to 35% from 2020-2030, which is higher than the average across all occupations. Despite this promising outlook, statistics coursework has been cited as challenging to students for several reasons, notably, a lack of interest (Aggarwal, 2018) and anxiety (Onwuegbuzie, 2004) when taking statistics courses. Furthermore, overall math anxiety in students internationally is significantly related to lower math assessment scores (Foley et al., 2017) and math anxiety has been reported to lead to avoidance of taking STEM classes and reduced grades in STEM courses, even when math ability was controlled for (Daker et al., 2021).

Statistics anxiety is a form of state anxiety marked by thoughts and feelings of fear, worry, and apprehension at the prospect of engaging with statistics related content and can result in delayed enrollment in statistics and research methods courses (Onwuegbuzie, 2004; Onwuegbuzie & Wilson, 2003; Zeidner, 1991), as well impact achievement potential in those courses. Statistics courses have been shown to produce some of the highest levels of anxiety for various demographics of university students (Bell, 2003; Chew & Dillon, 2014a; Onwuegbuzie et al., 2010). Students enrolled in human science and social science fields of study often prefer to learn applied human-interest content and avoid courses in research methodology and statistics courses, suggesting poor attitudes and low perceived relevance towards the latter disciplines (Aggarwal, 2018; Rajecki et al., 2005). Kaufmann and colleagues (2022) found that students who perceived statistics to be of low relevance had higher rates of statistics anxiety. The widespread prevalence of statistics anxiety and avoidance of taking statistics courses has contributed to the development of several instruments to study this phenomenon. These include such instruments as the Statistics Course Attitude Scale (Bendig & Hughes 1954), as well as the more recent Statistics Anxiety Rating Scale (STARS; Cruise et al., 1985) and Statistics Anxiety Scale (SAS; Vigil-Colet et al, 2008).

Statistics Anxiety Scale

The SAS is a 24-item inventory that was developed by Vigil-Colet, Lorenzo-Seva, and Condon (2008) in Spain. The items are Likert-type, ranging from 1 (no anxiety) to 5 (very much anxiety). The SAS was developed via two steps (Vigil-Colet et al., 2008). Firstly, feedback was gathered from statistics instructors with more than 10 years of experience. Secondly, some of the items were derived from the STARS, namely items 1-3, 6, 9-11, 14, 17, 18, and 22 (see Table 1). The participants consisted of 159 psychology students enrolled in a statistics course. Using exploratory factor analysis (EFA), the authors determined three subscales of statistics anxiety: asking for help anxiety (AHA), examination anxiety (EA), and interpretation anxiety (IA). Asking for help anxiety is defined as anxiety when asking an authority figure or peer questions about statistics. The items that loaded onto this factor were items 3, 5, 7, 12, 17, 21, 23, & 24. EA is defined as the experience of anxiety when taking a statistics examination. The items that loaded onto this factor were items 1, 4, 9, 11, 13, 14, 15, & 20. Lastly, IA is defined as the experience of anxiety when having to interpret data and understand formulae in the statistical context. The items that loaded onto this factor were items 2, 6, 8, 9, 10, 16, 18, & 22. The total variance explained by the factors was 70.9 %. The SAS and its constructs were also found to be internally consistent, with alpha coefficients of .924, .874, .819, and .911 respectively for AA, EA, IA, and the SAS. Factor correlations were moderate, ranging from .308 to .445. The SAS and its subscales also shared weak to moderate relationships with two other instruments: the trait subscale of the State Trait Anxiety Inventory (Spielberger et al., 1988), and the neuroticism component of the Eysenck Personality Question Revised (Eysenck et al., 1985), thus addressing convergent validity of the SAS. All three factors have a potential range of 8 - 40.

Chiesi and colleagues (2011) later confirmed this factor structure using Italian (n = 512) and Spanish (n = 336) participants. Using confirmatory factor analysis (CFA), they reported a reasonable fit with Italian participants, $\chi^2(248) = 668.361$, $p < .01$; Comparative Fit Index (CFI) = .926, Root Mean Square Error of Approximation (RMSEA) = .058, and Spanish participants, $\chi^2(246) = 556.296$, CFI = .926, RMSEA = .061. The factor models were considered to be equivalent across both countries, albeit with modifications. Similar to Chiesi et al. (2011), Chew and Dillon (2014b) found a modified three-factor model amongst a sample of students from Singapore and

Australia, χ^2(240] = 532.73, CFI = .92, Parsimony Goodness of Fit Index (PGFI) = .66 , RMSEA = .08 (90 % CI: .07, .09). In contrast, O'Bryant and colleagues (2021) found a modified two-factor model amongst American university students that did not differ from the three-factor model, $\chi^2(240) = 49.37$, p = .105, CFI = .959, RMSEA = .076, Standardized Root Mean Square Residual (SMSR) = .035. In this model, items 24 and 15 were excluded from AHA and EA respectively. Additionally, IA was not retained for this modified version. Overall, the three-factor model was reported to have a poor fit, with the two-factor model having a more acceptable fit.

Effects of Age, Gender, Format, and Race/Ethnicity

Stereotype threat is defined as anxiety resulting from the perception that the behavior of members of stigmatized groups may confirm negative stereotypes about those groups. A common stereotype threat related to statistics anxiety results from the negative stereotype of minorities and women being intellectually inferior (Beasley & Fischer, 2012; Shapiro & Williams, 2011; Steele, 1997), and is implicated in academic decline and attrition for these populations (Beasley & Fischer, 2012; Davis et al., 2006). Additionally, it has been demonstrated that "age non-traditional students" (i.e, age 25 and older) have more statistics anxiety, than traditional students (Bell, 2003), and "classroom traditional learners" have reported slightly less AHA versus online learners (Frey-Clark et al., 2019).

The main purpose of this study was to validate a modified version of the SAS for use with college students taking statistics courses taught at American universities in different formats (face-to-face vs. hybrid vs. online). The version of the SAS presented here had four major modifications: modifications of existing items to be more consistent with language used by students in the United States; modification of items to reflect shifts toward online education, especially during the COVID-19 pandemic; addition of items that address the use of statistical programs and packages; and items addressing peer mentoring. The secondary purpose was to explore what demographic differences (age, gender, format, race/ethnicity) were found in statistics anxiety. We hypothesized the following:

Hypothesis 1: items 3, 5, 7, 12, 17, 21, 23, & 24 would load onto the AHA factor; that items 1, 4, 9, 11, 13, 14, 15, & 20 would load onto EA; and that items 2, 6, 8, 9, 10, 16, 18, & 22 would load onto IA. This conforms to the original factor structure.

Alternatively, **Hypothesis 1a:** items 3, 5, 7, 12, 17, 21, & 23 would load onto AHA; and items 1, 4, 9, 11, 13, 14, & 20 would load onto EA. This conforms to the factor structure confirmed by O'Bryant et al. (2021).

Method

Materials

Content and Face Validity

This study was approved by the Institutional Review Board (2021-078-OI) at the authors' institution. The nominal group technique (NGT; Delbecq et al., 1975) was used to reach a consensus on items that needed to be reworded or adjusted, as well as for new survey item elicitation (Harb et al., 2021). The original scale developed by Vigil-Colet et al. (2008) was used for this process. In the Fall semester of 2020, the SAS items were reviewed by a panel of five university faculty who had experience either teaching statistics and/or who were heavily involved in quantitative methodology. These included faculty with experience teaching either undergraduate or graduate statistics courses, one survey methodologist, and one other faculty member regularly involved with quantitative-directed learning in the field of physiology. All items were first reviewed by each faculty member independently and suggestions for modifications were sent to the researcher, who moderated the process. The moderator had extensive qualitative research experience. All faculty members then met with the moderator to discuss modifications and additional survey items after initial feedback had been collated. Multiple rounds of discussion and voting took place regarding all adjusted and new survey items until a consensus was reached. The outcome of this meeting included changing the term "teacher" to "instructor", and the term "private teacher" to "tutor" throughout the instrument. There were also other modifications to items 3, 4, 9, 11, 12, 13, 15, 16, 17, 19, 20, and 23. Generally speaking, these modifications were meant to better represent the experiences of students taking classes in any format (i.e., traditional, online, or hybrid formats). For example, item 23, "going to the teacher's office to ask questions" was modified to "meeting with the instructor to ask questions," as those taking the class in an online format may not have the time or means to physically visit an instructor's office.

Items 25-28 were added, as many courses now offer a laboratory component whereby students must learn to operate one or more statistical programs ("packages"), which was not assessed in the original SAS. Items 29and 30 were added as it has been shown that modes of peer mentoring are effective learning tools for students in the STEM fields. For example, supplemental instruction has been shown to have favorable learning and socialization outcomes (Hurley et al., 2006). Gorvine and Smith (2015) have used creative (and less expensive) teaching methods to implement student collaboration in statistics, in which students experienced lower statistics anxiety, and had a higher preference for working in groups. These resulted in more favorable learning outcomes.

After the survey items were updated, the NGT was repeated with a group of seven students (Fall 2020) who already took an applied statistics course designed for allied health. This group represented the demographic from which we sampled our participants. It was completed for clarity and was therefore an additional step in face validation. Comments were made about adding items related to COVID-19, and general mathematics anxiety. These items were considered but not added, as the COVID-19 item was specific to the pandemic, and the SAS is specific to statistics, but not generalized to mathematics. Students reported satisfaction with all other items, and no other modifications were made after this step. Therefore, the final survey consisted of 30 Likert-type items ranging from 1 (no anxiety) to 5 (very much anxiety). The wording for each item may be found in Table 1. We also hypothesized the following: Hypothesis 2: Items 25-28 would load onto IA, as they represent levels of statistical analysis and interpretation. Hypothesis 3: Items 29 and 30 would load onto AHA, as they involved solicitation of assistance in solving statistical problems.

Wise's Attitude Toward Statistics Scale

Wise's Attitude Toward Statistics scale (ATS; Wise, 1985) was used in order to establish discriminant validity with the newly worded SAS. It is a 29-item instrument with a 5-point Likert scale. Of the 29 items, 14 are reverse scored. The items may be summed for two factors/subscales (Attitude toward Field and Attitude toward Course) or for a total score. Higher scores indicate more positive attitudes toward statistics, and the total score was chosen for this study.

Once this step was completed, students were recruited from several American universities representing states such as Texas, Colorado, and Illinois (Spring 2021 through Fall 2022). A link was sent either directly to instructors, or to chairpersons of departments in which statistics courses were taught. Chairpersons distributed the link to their respective instructors, who in turn shared the link with their respective students. Participants were not required to participate. The departments represented faculties of mathematics and the health sciences. At least three instructors were represented, though the total number is unknown, due to student anonymity. Students were anonymized as some of the students belonged to a class taught by one of the investigators, and anonymity was a step to minimize coercion. As students were anonymized, it was not possible to determine which instructors shared the link, or which classes/students participated for the instructors that did. Students were asked to complete the survey ≤ 2 days before or after an examination to take advantage of salient feelings of anxiety. Dates were requested from instructors and/or heads of department to maximize participation. Surveys were distributed using the Qualtrics electronic survey platform (Provo, Utah).

Participants

The data were collected during Fall 2021, Spring 2022, and Fall 2022 semesters. Participants (n = 352) were sampled from statistics courses from universities in Colorado, Illinois, and Texas. Participants had a mean age of 19.4 ± 3.4 years (n = 346), ranging from 17 to 60 years old, with six non-disclosing. Seventy-two students identified as Asian-American, 57 Black, two first-nations, 59 Hispanic/Latinx, 1 Pacific Islander, 154 White, with seven non-disclosing. There were five gender non-conforming, 250 females, and 96 males, with one student non-disclosing. Student majors varied across several disciplines. Some of the more represented majors included business-related (n = 42), psychology (n = 41), political science (n = 19), animal sciences (n = 17), communications (n = 8), kinesiology/exercise science (n = 17), (pre-) nursing (n = 16), community health/wellness promotion (n = 13), education (n = 8), and English (n = 7). Except for one of the education majors that declared a double major, no other students reported being a mathematics major. There were also 34 undeclared majors. One hundred and fifty-four students reported taking the course in a traditional (face-to-face) format, 161 took the course in an online format, and 37 students took the course in a hybrid (mix of traditional/online) format. Though taught at different institutions, the classes emphasized the application of statistical thinking and tools to problem solving. Topics typically found in such classes include probability, study design, correlation, linear regression, *z*-tests, *t*-tests, and analysis of variance.

Data Analysis

Data were analyzed using IBM SPSS version 28 (Armonk, NY). Descriptive statistics were obtained for all SAS items (1-30). Items 16 and 20 had skewness values of 1 and -1.1 respectively, while items 3 and 14 had kurtosis values of -1 and -1.1 respectively. All other items had skewness and kurtosis values between 1 and -1. Hence the data were treated as normal for the analyses presented in this study. Two exploratory factor analyses using principal axis factoring (EFA) were used to analyze the data, with one being used to analyze the items for the original 24-item scale (SAS-O), and the other for the modified 30-item scale (SAS-M). The purpose of this step was to address construct validity. The direct oblimin rotation method was used for both analyses, as subscales of statistics anxiety are shown to be related (Chew & Dillon, 2014b).

We interpreted EFA for several reasons. Firstly, the study by O'Bryant and colleagues (2021) used confirmatory factor analysis (CFA) and found a revised two-factor model with American students that contradicted the three-factor model found or confirmed in other studies. With two different models, there lies the possibility that the results found in Spanish (Vigil-Colet et al., 2008; Chiesi et al., 2011), Italian (Chiesi et al., 2011), Australian, and Singaporean (Chew & Dillon, 2014) students might not be generalized to American students. This raises a question of cross-cultural construct validity (likely in terms of the educational framework) of the SAS with the latter population. Furthermore, as the three-model factor was found to be a poor fit via the use of CFA, EFA can be used to re-examine these models (Schmitt, 2011).

We also modified and added items to the SAS. Several of these modifications allowed for the dynamic nature of instruction (e.g., advances in online teaching, increased availability of statistical programs, etc.) that have happened since the SAS' inception in 2008. We wished to explore the factor structure in order to see if items would remain on the same factors, or load onto different ones. We also wished to compare and report the factor structures and the amounts of variance explained by the factors found in SAS-O and SAS-M. Examples of this step's importance include the enhanced decision-making ability of instructors to choose a version for course evaluation, and enhanced decision-making for other scientists wishing to expand upon this work (e.g., replication with other EFA's as well as CFA).

In summary, we wished to explore the factor structure in our sample of American students without forcing the factors through the use of CFA for two reasons. Firstly, we did not wish to use CFA as an exploratory tool and force a poor-fitting three-factor model in American students. Secondly, we substantially modified the wording of the instrument from its original form. Cronbach's α and McDonald's ω analyses were used as follow-up analyses to determine the internal consistency of each factor/subscale and for ATS.

Cluster analysis is a family of exploratory methods used to dissect datasets into unique subgroups using relevant variables (Meyers et al., 2017). Two-step cluster analysis is a clustering algorithm that considers categorical and continuous variables (IBM Support, 2020). This method was used as a follow-up analysis to explore any potential differences in the five subscales of statistics anxiety between six categories of race/ethnicity, three categories of gender, and three teaching formats. This method was chosen instead of a more hypothesis-driven method (e.g., analysis of variance), as several of these categories differed in size (e.g., the low numbers of Pacific Islander, First Nations, and Gender Non-conforming participants). Cluster analysis thus offered the option to analyze without collapsing these small samples into larger groups, thus minimizing loss of information. However, the disadvantage of using this method is a potential lack of reproducibility in future studies. Follow-up chi-square tests of independence were used to determine the strength of association between cluster membership and the demographic variables (format, gender, race/ethnicity), while follow-up analyses of variance (ANOVA) were used to determine differences in cluster membership for age and each of the five subscales of statistics anxiety. The Bonferroni post hoc test was used for all ANOVA analyses except for application anxiety, where the Levene's Test for equal variances was violated. In this case, the Games-Howell post hoc test was used (for the variable of age). The items and their descriptive statistics are included in Table 1.

Table 1. Descriptive Statistics of the SAS Items

Item	**Mean ± S.D.**
1 Studying for an examination in a statistics course	3.1 ± 1.1
2 Interpreting the meaning of a table in a journal article	2.5 ± 1.0
3 Asking my statistics instructor for help with material I am having difficulty understanding.	2.5 ± 1.3
4 Realizing that I cannot do some problems that I thought were going to be easy when preparing for an upcoming exam.	3.7 ± 1.1
5 Asking a tutor to explain a topic that I do not understand	2.5 ± 1.2
6 Reading a journal article that includes some statistical analyses	2.4 ± 1.1
7 Asking the instructor how to use a probability table	2.5 ± 1.2
8 Trying to understand a mathematical demonstration	2.7 ± 1.1
9 Completing the final exam in a statistics course	4.0 ± 1.1
10 Reading about public policy statistics such as data about the census, employment, public health, and consumer spending	2.3 ± 1.0
11 Getting ready to start a statistics exam	3.7 ± 1.2
12 Asking the instructor about how to do a statistics assignment	2.4 ± 1.2
13 On the day before you take a statistics exam, feeling as though you might get a particular problem that you do not know how to answer.	3.6 ± 1.2
14 Waking up on the day that you take a statistics exam	3.2 ± 1.2
15 Realizing just before you start the exam that you have not prepared for a particular statistics problem.	4.0 ± 1.1
16 Copying a mathematical demonstration while the instructor is explaining it	2.0 ± 1.1
17 Asking the instructor for help with understanding a statistics software output	2.4 ± 1.3
18 Trying to understand the odds in a lottery	2.4 ± 1.1
19 Listening to a classmate discuss the results of a statistics problem they solved	2.0 ± 1.1
20 Taking a statistics exam without having had enough time to study	4.2 ± 1.1
21 Asking the instructor for help when trying to interpret a results table	2.4 ± 1.2
22 Trying to understand the statistical analyses described in the abstract of a journal article	2.5 ± 1.1
23 Meeting one-on-one with the instructor to ask questions	2.7 ± 1.4
24 Asking a tutor to explain to me how to do an assignment.	2.4 ± 1.2
25 Learning a new statistics software program (JASP, SPSS, R, SAS, Mplus, etc.)	3.3 ± 1.3
26 Trying to apply a statistical test to solve a problem (i.e., correlation, regression, t-test, ANOVA, etc.)	3.1 ± 1.2
27 Learning a new computer coding language to solve statistical problems (e.g., Python, R)	3.5 ± 1.3
28 Learning how to use a calculator to solve statistical problems.	2.3 ± 1.2
29 Asking a peer or a classmate to explain a topic that I do not understand.	2.5 ± 1.2
30 Asking a peer or a classmate to explain to me how to do an assignment	2.4 ± 1.2

Results

Bartlett's Test of Sphericity (BTS) revealed the rejection of the null hypothesis of an identity matrix, i.e., no correlations between any of the variables in this study. For SAS-O, BTS is reported as follows: approximate $\chi^2(276) = 3413.9$, $p < .001$. For SAS-M, BTS is reported as approximate $\chi^2(435) = 4586.9$, $p < .001$. Pearson product moment correlations between SAS-O/SAS-M and the total score of Wise's Attitude Toward Statistics scale (1985) were used to establish discriminant validity of SAS-O and SAS-M. All items correlated ≥ .3 with at least one other item. Each item had individual Kaiser-Meyer-Olkin (KMO) values ≥ .85, with the overall KMO being .93. Items 2, 6, 8, and 16 had communalities below .5, but were warranted for Inclusion due to meeting the previous two criteria. Hence, EFA was interpreted for the scales.

Original SAS

For SAS-O, there were three factors, explaining 64.3 % of the variance in the data. The factor explaining the most variance contained items 3, 5, 7, 12, 17, 21, and 23. Consistent with previous studies, this factor was called Asking for Help Anxiety. This was followed by the factor containing items 1, 4, 9, 11, 13, 13, 14, 15, and 20 (Examination Anxiety). Lastly, items 2, 6, 8, 10, 16, 18, 19, and 22 loaded onto Interpretation Anxiety. The factor structure, variance explained, eigenvalues, reliability analyses (α and ω), and descriptive statistics for these factors may be found in Table 2. Note that both α and ω are included as there have been calls to use the latter in preference to the former (e.g., Dunn et al., 2013).

Table 2. Factor Structure of SAS-O

Item	AHA	EA	IA	Communality	KMO
21	.87			.74	.95
17	.87			.76	.97
24	.86			.78	.94
12	.84			.73	.96
3	.82			.71	.94
7	.80			.69	.96
23	.78			.63	.95
5	.78			.62	.95
15		.83		.65	.92
11		.82		.71	.95
14		.80		.58	.94
13		.78		.59	.95
9		.77		.58	.95
4		.76		.52	.94
20		.70		.53	.89
1		.65		.48	.95
10			.80	.51	.92

Item	AHA	EA	IA	Communality	KMO
18			.72	.53	.95
22			.72	.56	.94
6			.70	.50	.91
19			.68	.47	.95
2			.66	.53	.94
8			.59	.48	.97
16			.57	.36	.93
				Total variance explained	
% variance	42.2	12.8	9.3	64.3	
Eigenvalue	10.1	3.1	2.2	15.4	
Cronbach α	.95	.91	.88	**Instrument α:** .94	
McDonald's ω	.95	.91	.88	**Instrument ω:** .94	
Mean ± S.D.	19.9 ± 8.6	29.6 ± 7.2	18.8 ± 6.3		
Present Range	8-40	9-40	8-38		
Potential range	8-40	8-40	8-40		

Modified SAS

There were five factors explaining 68.4 % of the variance in the SAS-M. While AHA and EA were consistent with the analysis of SAS-O, IA was consistent with SAS-O, except for the addition of item 28. The first new subscale of Application Anxiety (AA) contained items 25, 26, and 27. The second new subscale of Peer Anxiety (PA) contained items 29 and 30. The factor structure of SAS-M violates recommendations set forth by several scholars concerning having too few items per factor (e.g., Izquierdo et al., 2014). However, we chose to move forward with interpreting SAS-M because of the practical relevance of these items, especially given the demographic differences found in AA, discussed further below. The factor structure, variance explained, eigenvalues, reliability analyses, and descriptive statistics for these factors may be found in Table 3.

Table 3. Factor Structure of SAS-M

Item	AHA	EA	IA	AA	PA	Communality	KMO
21	.87					.74	.95
17	.87					.76	.97
24	.86					.78	.94
12	.84					.73	.97
3	.83					.73	.94
7	.81					.70	.96
5	.78					.62	.96
23	.78					.63	.95
11		.82				.71	.95
15		.82				.70	.91

Item	AHA	EA	IA	AA	PA	Communality	KMO
14		.79				.56	.94
13		.76				.60	.95
9		.75				.59	.96
4		.75				.53	.95
20		.69				.55	.91
1		.66				.49	.96
10			.76			.49	.92
22			.70			.57	.96
6			.69			.51	.93
18			.68			.53	.95
2			.67			.54	.95
19			.66			.49	.96
28			.65			.48	.97
16			.59			.37	.94
8			.58			.50	.96
25				-.83		.82	.88
27				-.80		.73	.91
26				-.58		.68	.93
29					.96	.90	.86
30					.93	.84	.86
						Total variance explained	
% variance	41.0	11.2	8.3	4.4	3.5	68.4	
Eigenvalue	12.3	3.4	2.5	1.3	1.1	20.6	
Cronbach α	.95	.91	.89	.88	.94	**Instrument α:** .95	
McDonald's ω	.95	.91	.89	.89	-*	**Instrument ω:** .95	
Mean ± S.D.	19.9 ± 8.6	29.6 ± 7.2	21.1 ± 7.1	9.9 ± 3.3	4.8 ± 2.3		
Present Range	8-39	9-40	9-42	3-15	2-10		
Potential range	8-40	8-40	9-45	3-15	2-10		

*Note: McDonald's ω not provided, as it will not compute reliability for less than three items.

Discriminant Validity

The internal consistency for Wise's Attitude Toward Statistics scale (ATS; 1985) was .92 for both Cronbach's α and McDonald's ω. As shown in Table 4, the SAS and all subscales from both instruments were inversely related to the total score of ATS. All correlation null hypotheses were rejected at the $p = .01$ level.

Table 4. Relationships between the SAS Subscales and Wise's ATS

SAS-O						
	AHA	**EA**	**IAO**	**SAS-O**		
Asking for Help Anxiety (AHA)						
Examination Anxiety (EA)	.45					
Interpretation Anxiety (original; IAO)	.56	.47				
SAS-O	.85	.78	.81			
Attitude Toward Statistics Scale (ATS)	-.22	-.28	-.33	-.33		
SAS-M						
	AHA	**EA**	**IAN**	**AA**	**PA**	**SAS-M**
EA	.45					
Interpretation Anxiety (new; IAN)	.67	.48				
Application Anxiety (AA)	.37	.55	.58			
Peer Anxiety (PA)	.63	.37	.54	.39		
SAS-M	.83	.77	.83	.69	.69	
ATS	-.22	-.28	-.34	-.18	-.23	-.33

*H_0 rejected at the p = .01 level; **H_0 rejected at the p = .001 level

Two-Step Cluster Analysis

A total of 346 students were clustered. This is because six students did not provide their ages and were not included in this analysis. These students were still analyzed descriptively as a separate "unclassified" group (U). The Silhouette measure of cohesion and separation was classified as poor. Overall, the best predictors in order from highest to lowest priority were class format, gender, race/ethnicity, EA, age, AA, AHA, PA, and IA. Cluster 1 had 90 participants, and consisted of only males. Class format had the lowest predictor performance within this group. This cluster had significantly less AA and EA than all other groups and had significantly less AHA than Cluster 2. No other clusters were different in AHA, AA, and EA.

Cluster 2 appeared to be the most non-traditional group. It had 73 participants and was predominantly female. It had the lowest proportion of traditional learners, the highest number of students older than 25 (n = 8), and contained most of the Hispanic participants, as well as hybrid learners. This cluster was significantly older than Cluster 4, had the highest age range (17-60), contained the most gender diversity, and as aforementioned had higher AHA than Cluster 1.

Cluster 3 had 89 participants. It had the highest number of white participants, was predominantly female, and contained two gender non-conforming participants. Cluster 4 had 94 participants and was characterized by all female students taking a traditional course. Descriptively, this was also the youngest cluster and was significantly younger than Cluster 2. Generally, Clusters 3 and 4 were similar. However, race/ethnicity had a higher predictor performance for Cluster 3, while age had a higher predictor performance for Cluster 4. Differences in Cluster demographics may be found in Tables 5 and 6.

Table 5. Cluster Frequencies

Demographic	**Clusters**					**Group Totals**	**Exploratory χ^2 Test of Independence**
Race/Ethnicity	**1**	**2**	**3**	**4**	**U**ł		
Asian American	23	5	16	28	0	72	$\chi^2(18) = 141.9$, $p < .001$, V = .37
Black/African American	11	4	18	19	5	57	
American Indigenous	0	1	1	0	0	2	
Hispanic	17	42	0	0	0	59	
Pacific Islander	0	0	0	1	0	1	
White/Caucasian	35	20	54	44	1	154	
Prefer not to answer	4	1	0	2	0	7	
Format							
Traditional	38	22	89	0	5	154	$\chi^2(6) = 291.8$, $p < .001$, V = .65
Online only	44	22	0	94	1	161	
Hybrid	8	29	0	0	0	37	
Gender							
Male	90	3	0	0	3	96	$\chi^2(9) = 340.1$, $p < .001$, V = .57
Female	0	66	87	94	3	250	
Gender Non-conforming	0	3	2	0	0	5	
Prefer not to answer	0	1	0	0	0	1	
Cluster Totals	90	73	89	94	6	352	

ł not used in the analysis.

Table 6. Cluster Differences in the SAS

	1	2	3	4	Uł	Exploratory Analysis of Variance
Age	19.0 ± 1.0	21.0 ± 6.9***	19 ± 1.3	18.9 ± 1.2	-	$F(3, 342) = 7.6$, $p < .001$, $\eta^2_p = .06$
AHA	17.5 ± 7.5*	22.5 ± 8.6	19.9 ± 8.6	20.3 ± 8.8	18.8 ± 12.3	$F(3, 342) = 7.6$, $p < .01$, $\eta^2_p = .04$
EA	26.3 ± 7.8**	32.2 ± 6.0	29.8 ± 6.6	30.6 ± 6.6	28.8 ± 10.6	$F(3, 342) = 11$, $p < .001$, $\eta^2_p = .09$
IA	20.1 ± 7.2	21.9 ± 6.8	20.5 ± 6.3	21.9 ± 7.5	25.7 ± 11.3	$F(3, 342) = 1.6$, $p > .05$, $\eta^2_p = .01$
AA	8.8 ± 3.3**	10.8 ± 3.0	10.2 ± 3.2	10.1 ± 3.4	10.2 ± 4.0	$F(3, 342) = 5.7$, $p < .001$, $\eta^2_p = .05$
PA	4.4 ± 2.3	5.1 ± 2.1	5.0 ± 2.5	4.8 ± 2.3	6.3 ± 2.6	$F(3, 342) = 1.7$, $p > .05$, $\eta^2_p = .02$

*Cluster 1 different from Cluster 2 only; **Cluster 1 different from all other clusters; ***Cluster 2 different from Cluster 4 in age; ł not used in analysis.

Discussion

In this study, the SAS was modified to be used for American students, as well as for students who took the course in multiple formats (i.e., traditional, online, and hybrid formats). Items were also added to the modified version that pertained to peer mentoring and the use of statistical programs and packages. For the SAS-O, there was a three-factor solution, with AHA (Asking for Help Anxiety) explaining 42% of the variance, EA (Examination

Anxiety) explaining about 13 %, and IA (Interpretation Anxiety) explaining about 9 %, for a total of about 64 % of the variance. This three-factor structure is similar to the original (Vigil-Colet et al., 2008), in which 70 % of the variance was explained, therefore Hypothesis 1 exists in a sample of American students. Previous studies have shown the factor structure of the SAS to be consistent amongst Spanish, Italian (Chiesi et al., 2011), Singaporean, and Australian (Chew & Dillon, 2014b) psychology students. The three-dimensional factor structure of the original SAS is now shown to be consistent amongst students in the United States taking statistics courses across several institutions and several instructional formats.

In contrast, the amount of variance explained by each factor differed. Vigil-Colet and colleagues (2008) found that AHA accounted for about 28 % of the anxiety, with EA and IA accounting for about 14 % and 19 % of the variance respectively. Whereas the variance was more equally distributed in the Vigil-Colet study, asking for help accounted for more than twice the amount of variance than the other two factors in this study. The majority of students ranged between low and moderately high levels of AHA anxiety. Reasons for this variation are not clear but may include individual dynamics between the students and their respective instructors (such as availability) during the COVID-19 pandemic when these data were collected. However, it is a positive indicator that the mode was the lowest score on the subscale and that the frequency of this score is more than 1/10 of the entire sample. Expectedly, EA tended to be high, with a mean of about 30, and a mode of 35 ($n = 23$), followed by scores of 36, 34, and 31 ($n = 20$ for each), followed by 30, 37, and 38 ($n = 19$). Students tend to feel high anxiety around examination times when this survey was administered, and the higher scores of this subscale could also have been due to the timing of survey completion when feelings of anxiety were salient.

Interpretation anxiety is the only dimension that differed slightly between the original SAS and the modified SAS with new items added. For the modified version, item 28 was included. This item was stated as “Learning to use a calculator to solve statistical problems” contains items that address the pragmatic application of statistical knowledge (i.e., items 10, 22, 6, 18, 2), as well as mathematical interpretation and problem-solving (items 19, 16, 8). Given that the calculator is the tool commonly used to solve mathematical problems, it makes sense that item 28 loaded onto IA.

Items 25, 26, and 27 addressed the learning and application of programs or coding to analyze data. These items were included, because many statistics courses have a laboratory component that makes use of a graphical user interface (e.g., SPSS) or syntax-driven program (e.g., R) to screen and analyze datasets. This loaded onto a unique dimension, AA, which addresses anxiety related to the selection of statistical tests to analyze data and learning how to operate these programs. It should be noted that AA does contain an element of interpretation (item 27), but this pertains to interpreting the quality of data to be analyzed, as opposed to analyzing the results. This finding rejected our hypothesis that these three items would load onto IA. That said, AA was most correlated to IA, with both dimensions sharing about 33 % of variance (squared correlation coefficient in Table 4) with each other. This finding suggests the importance of being mindful when instructing the use of statistical programs.

Items 29 and 30 addressed asking for help, but from a peer, and loaded onto a fifth dimension, PA. It was separate from the AHA dimension, as AHA had to do with asking for help from an authority figure such as an instructor or tutor. This finding also rejected our hypothesis that these items would load onto AHA. This is important to know as students may process information differently from a peer than they do from an instructor or tutor, and the use of peer mentoring is shown to be an effective teaching strategy in statistics (Gorvine & Smith, 2015), as well as other STEM courses (Hurley et al., 2006).

Comparison to the Original SAS

As aforementioned, the factor structure of SAS-O is consistent with that found by Vigil-Colet et al. (2008). They reported that the SAS factor structure explained about 71 % of the common variance, with AHA, EA, and IA explaining about 28 %, 24 %, and 19 % of the variance respectively. In contrast, the factors explained a total of about 64 % of the variance, with AHA, EA, and IA explaining about 42%, 13 %, and 9 % respectively in this study.

Demographic Differences in Statistical Anxiety

As aforementioned, two-step cluster analysis created four groups of students that were primarily driven by demographics. Cluster 1 consisted of only male participants. Cluster 2 was mostly female and appeared to be the most non-traditional. Clusters 3 and 4 were somewhat similar, with the latter being younger females only. Cluster membership was most associated with format (42 % shared variance), followed by gender (32 %), and race/ethnicity (13.7 %). Cluster membership explained differences in EA (9 %, a medium-to-large effect, [Cohen, 1988]), AA (5 %), and AHA (4 %, both small-to-medium effects). There were no other differences in anxiety between clusters.

Cluster 1 had less AA and EA than all of the other clusters, and less AHA than Cluster 2. There were no other differences between clusters in terms of statistical anxiety. These findings suggest that males felt more empowered asking for instructor assistance, operating statistical programs, and when preparing to take examinations. Cluster 2 represented a unique intersection of students that were generally female non-traditional students and reported being slightly more anxious than Cluster 1 when asking for instructor assistance. However, they were no different from the other clusters for any statistical anxiety subscales.

Cluster 2's difference from Cluster 1 adds to the findings of previous research. Bell (2003) found that non-traditional students had significantly more Test and Class Anxiety than traditional students. They also trended higher in AHA and IA, though these differences were not statistically significant. Frey-Clark et al. (2019) similarly found that online students reported slightly more AHA than traditional peers. The unclassified group had the largest variation in all SAS subscales, and likely contained participants that potentially belonged to several clusters. Generally, there was a trend of female statistics students feeling slightly more anxious when it comes to the use of statistical programs and asking for help with examinations. Although the differences were not large, they were meaningful (see Table 6), which is consistent with previously reported trends in the literature for female students (Trassi et al., 2022). Given the challenges students experience with statistics anxiety, various strategies may be implemented to reduce these disparities.

Implications

Multi-faceted approaches may be useful in addressing students' anxiety in statistics. For example, Kiss and colleagues (2021) recommend including guided breathing exercises for students, expressive writing, and cognitive re-appraisal of their anxiety. Instructors are encouraged to foster a positive and supportive learning environment incorporating applied, pragmatic, and interactive elements of instruction. Instructors should show care toward student learning and promote positive rapport where students do not feel judged to seek help from instructors. Students should be encouraged to solve class problems as a team, rather than being isolated. By using teamwork to solve problems as part of a supportive learning community, students can validate one another's effort and progress, and their self-concept and confidence relative to statistical concepts may improve. Instructors may also help to alleviate examination anxiety by offering rough templates such as a mock examination beforehand (Kiss et al., 2021)

Kaufmann and colleagues (2022) evaluated statistics anxiety along with self-efficacy for statistical knowledge with a sample of university psychology students and found self-efficacy to have a moderating effect on statistics performance. The practical recommendations from Kiss and colleagues (2021) to build a positive learning community in the university classroom that incorporates pragmatic and applied teaching examples and encourages peers to work together and support one another in their learning process of mastering statistical competencies, complement the findings of Kaufmann and colleagues (2022). These pedagogical strategies directly support students in building self-efficacy (Bandura, 1994) through peer role modeling in group learning and positive verbal persuasion from the instructor and peers, while making the topic of statistics more relevant to students through practical and applied teaching examples. These recommendations may help to collectively reduce students' statistics anxiety.

Interventions such as supplemental instruction (SI) have also been shown to improve outcomes in performance in STEM disciplines by using student assistants and group work to clarify concepts (Arendale, 2006). Used as an attrition prevention method in high-risk courses, it employs a SI supervisor who identifies high-risk courses and manages SI peer leaders (Hurley et al., 2006). The SI supervisor is also responsible for enlisting faculty support and evaluating the SI program as well as individual sessions (Hurley et al., 2006). The SI peer leaders are students who have demonstrated mastery in the course and attended SI training sessions. The SI peer leaders model effective study strategies for students, such as the setting of short-term goals, whereby passing tests and the course are broken down into smaller, more manageable tasks. The SI peer leaders also play crucial roles in the cognitive components of the model by helping students learn to process new experiences and information via socialization and teamwork. They mediate student communication and conflict about course content and teach students to study independently while working together to achieve common goals, and improve self-efficacy. The SI peer leaders serve as positive role models and help to expose students to positive experiences to difficult courses, thereby helping additionally enhancing self-efficacy. While having several advantages, SI can be expensive and requires many staff to be trained (Dawson et al., 2014).

While SI may be too resource intensive to be broadly implemented, individual instructors can use creative methods or choose from the variety of pedagogical strategies aforementioned to reduce performance gaps and anxiety among statistics students. Models such as the flipped classroom approach may be used, whereby group assignments may be implemented in statistics courses to improve self-efficacy. The flipped classroom is defined as a system whereby the majority of teacher-centered instruction is removed from the classroom in favor of student-centered learning activities (Låg & Sæle, 2019). Students can be encouraged to study concepts beforehand, maximizing the use of active and applied learning with statistical programs and the completion of assignments in class. Instructors can play a role in mediating student group discussions around difficult material, thereby teaching students to work together to solve problems and improve confidence. For example, Gorvine and Smith (2015) found that students felt less anxiety, preferred group work, and had better learning outcomes using similar approaches in a statistics course for psychology students. For fully online courses, instructors may consider incentivizing queries and responses around difficult material on discussion boards to encourage students to socialize and communicate with one another about problematic material, with the instructor clarifying concepts and engaging in the discussions as needed.

Conclusion

The findings of this study provide a modified version of the SAS for use among college students in the United States which shows a consistent and reliable factor structure. The three-dimensional factor structure of the original SAS explained about 64 % of the common variance in a sample of 352 students. This modified version of the SAS included six items addressing the use of statistical programs and asking a peer for help. A total of five factors explained 68 % of the common variance in the sample. The factors were correlated to one another, negatively correlated with Wise's Attitude Towards Statistics scale, and were internally consistent using Cronbach α reliability analysis. Groups of female and non-traditional students felt more anxiety around examinations, and the use of statistical programs. Whereas the group of non-traditional students felt more anxiety when asking for help when compared to male students. This scale could be useful in assessing students' statistics anxiety across dimensions and used as a pre-post assessment tool to assess the effectiveness of some of the suggested pedagogical strategies discussed here to decrease students' statistics anxiety and increase academic performance and persistence in statistics and other research methods and STEM related courses.

Limitations and Recommendations

There were several limitations to this study. Firstly, this study took place during the COVID-19 pandemic. While the deleterious effects of the pandemic on mental health are well documented (Lee et al., 2021), any potential effect on responses to this instrument are neither known nor understood. While we have presented a rationale as to why we created six new items, the resulting two new dimensions explained only an additional 4 % of variance (i.e., about two items). Additionally, the factors created also violate recommendations set forth for the number of items per factor. However, pragmatic reasons exist for the inclusion of these items.

While differences were seen between males and females in EA and AA, we did not collect qualitative data to fully understand this phenomenon. While no differences were seen in PA, replication is important to verify these findings. As cluster analysis is an unsupervised method, the results found in that portion of the study may be difficult to replicate. We also did not collect grades to establish relationships with student performance and anxiety, or to add as an outcome for cluster analysis. While not a current objective, it is a good potential follow-up study. Other unsupervised methods such as latent profile analysis may also be useful in visualizing hidden groups of students, and replication of studies establishing relationships with personality and the SAS-M may be insightful in determining personality traits that predict AA and PA. Other recommendations for future study include the use of confirmatory factor analysis with different samples to confirm the factor structure of the SAS-O and the SAS-M.

Acknowledgements

We express gratitude toward the student participants recruited for this study, as well as for all edits and suggestions that contributed to the quality of this paper.

References

Aggarwal, R. (2018). Statistical literacy for healthcare professionals: Why is it important? *Annals of Cardiac Anaesthesia, 21*(4), 349-350. https://doi.org/10.4103/aca.ACA_177_18

Arendale, D. (2006). Understanding the supplemental instruction model. *New Directions for Teaching and Learning, 60,* 11-21.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). Academic Press.

Beasley, M. A., & Fischer, M. J. (2012). Why they leave: the impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education, 15*(4), 427-448. https://doi.org/10.1007/s11218-012-9185-3

Beck, T. (2013). The importance of a priori sample size estimation in strength and conditioning research. *Journal of Strength and Conditioning Research*, *27*(8), 2323-2337.

Bell, J. A. (2003). Statistics anxiety: the nontraditional student. *Education*, *124*(1), 157+. https://link.gale.com/apps/doc/A108911218/AONE?u=googlescholar&sid=bookmark-AONE&xid=75438a14

Bendig, A. W., & Hughes, J. B. (1954). Student attitude and achievement in a course in introductory Statistics. *Journal of Educational Psychology*, *45*(5), 268–276. https://doi.org/10.1037/h0057391

Bureau of Labor Statistics, U.S. Department of Labor (2022). Occupational Outlook Handbook, Mathematicians and Statisticians. Retrieved from: https://www.bls.gov/ooh/math/mathematicians-and-statisticians.htm

Chew, P. K. H., & Dillon, D. B. (2014a). Statistics Anxiety Update: Refining the Construct and Recommendations for a New Research Agenda. *Perspectives on Psychological Science, 9*(2), 196-208.

Chew, P. K. H., & Dillon, D. B. (2014b). Reliability and validity of the Statistical Anxiety Scale among students in Singapore and Australia. *Journal of Tropical Psychology, 4*(7), 1-7. https://doi.org/10.1017/jtp.2014.7

Chiesi, F., Primi, C., & Carmona, J. (2011). Measuring statistics anxiety: Cross-country validity of the Statistics Anxiety Scale (SAS). *Journal of Psychoeducational Assessment, 29*, 559-569.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985, August). Development and validation of an instrument to measure statistical anxiety. In *American Statistical Association Proceedings of the Section on Statistical Education* (Vol. 4, No. 3, pp. 92-97). American Statistical Association.

Daker, R. J., Gattas, S. U., Sokolowski, H. M., Green, A. E., & Lyons, I. M. (2021). First-year students' math anxiety predicts STEM avoidance and underperformance throughout university, independently of math ability. *Science of Learning, 6*(1), 17. https://doi.org/10.1038/s41539-021-00095-7

Davis, C., Aronson, J., & Salinas, M. (2006). Shades of Threat: Racial Identity as a Moderator of Stereotype Threat. *Journal of Black Psychology, 32*(4), 399–417. https://doi.org/10.1177/0095798406292464

Dawson, P., van der Meer, J., Skalicky, J., & Cowley, K. (2014). On the Effectiveness of Supplemental Instruction: A Systematic Review of Supplemental Instruction and Peer-Assisted Study Sessions Literature Between 2001 and 2010. *Review of Educational Research, 84*(4), 609-639. https://doi.org/10.3102/0034654314540007

Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes.* Scott, Foresman.

Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Epstein I., Santa Mina, E., Gaudet, J., Singh, M.D., & Gula, T. (2011). Teaching statistics to undergraduate nursing students: An integrative review to inform our pedagogy. *International Journal of Nursing Education Scholarship,* 8(1). https://doi.org/10.2202/1548-923X.2234

Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences, 36,* 1907-1920.

Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The math anxiety-performance link: A global phenomenon. *Current Directions in Psychological Science, 26*(1), 52-58. https://doi.org/10.1177/0963721416672463

Frey-Clark, M., Natesan, P., & O'Bryant, M. (2019). Assessing Statistical Anxiety Among Online and Traditional Students. *Frontiers in psychology*, *10*, 1440.https://doi.org/10.3389/fpsyg.2019.01440

Gorvine, B.J. & Smith, H.D. (2015). Predicting student success in a psychological statistics course emphasizing collaborative learning. *Teaching of Psychology, 42*(1), 56-59.

Hanoch, Y., & Pachur, T. (2004). Nurses as information providers: Facilitating understanding and communication of statistical information. *Nurse Education Today, 24*(3), 236-243.

Harb, S. I., Tao, L., Peláez, S., Boruff, J., Rice, D. B., & Shrier, I. (2021). Methodological options of the nominal group technique for survey item elicitation in health research: A scoping review. *Journal of Clinical Epidemiology, 139*, 140-148. https://doi.org/10.1016/j.jclinepi.2021.08.008

Hartland, G. V. (2020). Statistical analysis of physical chemistry data: Errors are not mistakes. *Journal of Physical Chemistry B, 124*(11), 2061-2064. https://doi.org/1021/acs.jpcb.0c01404

Hayat, M. J., Powell, A., Johnson, T., & Cadwell, B. L. (2017) Statistical methods used in the public health literature and implications for training of public health professionals. *PLoS ONE 12*(6), e0179032.

https://doi.org/10.1371/journal.pone.0179032

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2008). Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise 41*(1), 3-12. https://doi.org/10.1249/MSS.0b013e31818cb278

Hurley, M., Jacobs, G., & Gilbert, M. (2006). The basic SI model. *New Directions for Teaching and Learning, 106,* 11-22. https://doi.org/10.1002/tl.229

IBM Support (2020). *Clustering binary data with K-Means (should be avoided).* Retrieved from: https://www.ibm.com/support/pages/clustering-binary-data-k-means-should-be-avoided

Izquierdo, I., Olea, J., & Abad, F. J. (2014). *Exploratory factor analysis in validation studies: uses and recommendations.* https://doi.org/10.7334/psicothema2013.349

Kaufmann, L., Ninaus, M., Weiss, E. M., Gruber, W., & Wood, G. (2022). Self-efficacy matters: Influence of students' perceived self-efficacy on statistics anxiety. *Annals of the New York Academy of Sciences*, *1514*(1). https://doi.org/10.1111/nyas.14797

Kiss, A., Harari, R., & Vukovic R. (2021). *Reducing the Impact of Statistics Anxiety in College Classrooms | The Accessible Toolbox.* https://commons.mtholyoke.edu/theaccessibletoolbox/teaching-strategies-student-perspectives/math-anxiety/strategies/

Låg, T., & Sæle, R. G. (2019). Does the Flipped Classroom Improve Student Learning and Satisfaction? A Systematic Review and Meta-Analysis. *AERA Open*, *5*(3). https://doi.org/10.1177/2332858419870489

Lee, J., Solomon, M., Stead, T., Kwon, B., & Ganti L. (2021). Impact of COVID-19 on the mental health of US college students. *BMC Psychology, 9*(95), 1-10. https://doi.org/10.1186/s40359-021-00598-3

Meyers, L., Gamst, G., & Guarino, A. (2017). *Applied Multivariate Research* (2nd ed.). SAGE, Thousand Oaks, CA.

O'Bryant, M., Batley, P.N., & Onwuegbuzie, A.J. (2021). Validation of an adapted version of the Statistical Anxiety Scale in English and its relationship to attitudes toward statistics. *SAGE Open*, *11*(1). https://doi.org/10.1177/21582440211001378

Onwuegbuzie, A. J. (2004). Academic procrastination and statistics anxiety. Assessment & Evaluation in Higher Education, 29(1), 3–19. https://doi.org/10.1080/0260293042000160384

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics Anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in higher education, 8*(2), 195-209.

Onwuegbuzie A. J., Leech N. L., Murtonen M., Tähtinen J. (2010). Utilizing mixed methods in teaching environments to reduce statistics anxiety. *International Journal of Multiple Research Approaches, 4, 28–39*. doi: 10.5172/mra.2010.4.1.028

Pfaff, T. J., Sipos, M., Sullivan, M. C., Thompson, B.G., & Tran, M.M. (2013). The use of statistics in experimental physics. *Mathematics Magazine, 86*(2), 120-131. https://doi.org/10.4169/math.mag.86.2.120

Pollard, D. A, Pollard, T. D., & Pollard, K. S. (2019). Empowering statistical methods for cellular and molecular biologists. *Molecular Biology of the Cell, 30*(12), 1359-1368. https://doi.org/10.1091/mbc.E15-02-0076

Rajecki, D. W., Appleby, D., Williams, C. C., Johnson, K., & Jeschke, M. P. (2005). Statistics can wait: Career plans activity and course preferences of American psychology undergraduates. *Psychology Learning & Teaching, 4*(2), 83–89. https://doi.org/10.2304/plat.2004.4.2.83

Schmitt, T. A. (2011). Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment, 29*(4), 304–321. https://doi.org/10.1177/0734282911406653

Shapiro, J. R., Williams, A. M. (2011). The Role of Stereotype Threats in Undermining Girls' and Women's Performance and Interest in STEM Fields. *Sex Roles* **66**, 175–183. https://doi.org/10.1007/s11199-011-0051-0

Spielberger, C. D., Gorsuch, R. L., & Lushese, R. E. (1988). State Trait Anxiety Inventory. Madrid: TEA Ediciones S.A.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Trassi, A. P., Leonard, S. J., Rodrigues, L. D., Rodas, J. A., & Santos, F. H. (2022). Mediating factors of statistics anxiety in university students: a systematic review and meta-analysis. Annals of the New York Academy of Sciences, 1512(1), 76-97. https://doi.org/10.1111/nyas.14746

Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L. (2008). Development and validation of the Statistical Anxiety Scale. *Psicotherma, 20*(1), 174-180. https://doi.org/10.1037/t62688-000

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement, 45*(2),401-405. https://doi.org/10.1177/00131644850450022

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British journal of educational psychology, 61*(3), 319-328. https://doi.org/10.1111/j.2044-8279.1991.tb00989.x

Zhan, W., Fink, R., & Fang, A. (2010). Application of statistics in engineering technology programs. *American Journal of Engineering Education, 1*(1), 65-78. https://doi.org/10.19030/ajee.v1i1.793

Author Information

Keston G. Lindsay

http://orcid.org/0000-0002-2039-7477

University of Colorado Colorado Springs

United States

Contact e-mail: *klindsay@uccs.edu*

Jessica B. Kirby

http://orcid.org/0000-0003-0063-6546

University of Colorado Colorado Springs

United States

Brynn C. Adamson

http://orcid.org/0000-0002-1796-4047

University of Colorado Colorado Springs

United States