

# USING ATTITUDES AND ANXIETIES TO PREDICT END-OF-COURSE OUTCOMES IN ONLINE AND FACE-TO-FACE INTRODUCTORY STATISTICS COURSES

WHITNEY ALICIA ZIMMERMAN  
*The Pennsylvania State University*  
*waz107@psu.edu*

STEFANIE R AUSTIN  
*The Pennsylvania State University*  
*srb5302@psu.edu*

## ABSTRACT

*An abbreviated form of the Statistics Anxiety Rating Scale (STARS) was administered to online and face-to-face introductory statistics students. Subscale scores were used to predict final exam grades and successful course completion. In predicting final exam scores, self-concept, and worth of statistics were found to be statistically significant with no significant difference by campus (online versus face-to-face). Logistic regression and random forests were used to predict successful course completion, with campus being the only significant predictor in the logistic model and face-to-face students being more likely to successfully complete the course. The random forest model indicated that self-concept and test anxiety were overall the best predictors, whereas separately test anxiety was the best predictor in the online group and self-concept was the best predictor in the face-to-face group.*

**Keywords:** *Statistics education research; Course completion; Online education; Statistics attitudes; Statistics anxiety*

## 1. INTRODUCTION

Attitudes and anxiety have been studied in the context of postsecondary statistics education and the results suggest that these are constructs that may impact students' abilities to perform in such a course (e.g., Malik, 2015; Onwuegbuzie, 2004; Williams, 2013, 2015; Zeidner, 1991). However, with a handful of exceptions (e.g., DeVaney, 2016; Gundlach, Richards, Nelson, & Levesque-Bristol, 2015; Suanpang, Petocz, & Kalceff, 2004; Zimmerman & Johnson, 2017), the majority of studies have focused on students enrolled in face-to-face statistics courses. With increasing numbers of students enrolling in online courses (Allen & Seaman, 2017), these constructs should be examined in the online learning context as well.

An online course is one in which all or nearly all instruction is delivered via the Internet. This is in contrast to traditional face-to-face courses or hybrid courses, the latter which combine online and face-to-face instruction. According to Allen and Seaman's 2017 report on the current status of online education in the United States, 14.3% of all students in higher education are exclusively online students. In addition to those students, another 15.4% of all students in higher education are taking a combination of online and face-to-face courses. In total, more than 6 million students are taking at least one online course. There was a 11.0% overall increase in online enrollments from fall 2012 to fall 2015. Although there was a decrease in online enrollments at for-profit institutions during that time frame, there were larger increases in online enrollments at not-for-profit and public institutions.

The present study compares how statistics attitudes and anxieties relate to student success in face-to-face sections of a course versus in online sections of the same course. Through this observational study, the relations between attitudes and anxieties and performance on final exams are examined.

Additionally, the ability of these variables to predict successful course completion is compared for face-to-face and online students.

## 2. REVIEW OF LITERATURE

We reviewed literature concerning attitudes and anxieties in relation to academic performance in the context of postsecondary statistics courses. Research tends to show positive relationships between attitudes and academic performance and negative relationships between anxiety and academic performance. Whereas most research on statistics anxiety and attitudes has focused on face-to-face students, a few studies involving online students do exist. Below we summarize our review of literature concerning anxiety and attitudes in statistics courses, followed by a comparison of online and face-to-face students.

### 2.1. ATTITUDES AND ANXIETY TOWARD STATISTICS

An attitude is defined by the Dictionary of Statistics & Methodology as, “A positive or negative evaluation of and disposition toward persons, groups, policies, or other objects of attention” (Vogt, 2005). Social psychologists have been studying attitudes for many years because they may be used to predict behaviors (Crano & Prislin, 2006). The construct has also become popular amongst statistics education researchers in recent years (e.g., Vanhoof et al., 2011; see also Volume 11, Number 2, a special issue on research on statistics attitudes, in the *Statistics Education Research Journal*).

In addition to learning the content knowledge that is presented in a statistics course, students’ attitudes concerning statistics are also very important. Statistics instructors often believe that students should be able to see the value in learning statistics and should be confident in their abilities to apply what they have learned in their statistics course in real life in order to be considered “statistically literate” (Pearl, et al., 2012; Ramirez, Schau, & Emmioğlu, 2012; Schau, Millar, & Petocz, 2012).

Emmioğlu and Capa-Aydin (2012) conducted a meta-analysis summarizing studies that examined statistics attitudes using the Survey of Attitudes Towards Statistics (SATS-24; see Schau, Stevens, Dauphinee, & Del Vecchio, 1995) which originally consisted for four subscales: cognitive competence, affect, value, and difficulty. Measures of achievement included course and exam grades. They reviewed 17 studies and computed a 95% confidence interval for the correlations between achievement and each of the four SATS subscales. The strongest correlations were between achievement and affect (95% CI of  $\rho = (0.28, 0.32)$ ) and achievement and cognitive competence (95% CI of  $\rho = (0.28, 0.32)$ ). Correlations were slightly lower for achievement and value (95% CI of  $\rho = (0.19, 0.23)$ ) and achievement and difficulty (95% CI of  $\rho = (0.17, 0.22)$ ). These results suggest that there are moderately strong correlations between attitudes and achievement in postsecondary statistics courses.

Whereas the majority of studies have examined attitudes in students enrolled in face-to-face statistics courses, Gundlach et al. (2015) did compare undergraduate students enrolled in web-augmented traditional sections, online sections, and flipped sections of an introductory course taught by the same instructor. They administered the six-subscale version of the SATS (*SATS-36*; see Schau, 2003) at the beginning and end of a semester. There were significant time (pretest/posttest) by course section (traditional, online, flipped) interactions for the affect and perceived easiness subscales. Comparing the traditional and online sections, for both affect and perceived easiness, ratings were higher for the traditional group at the beginning of the course and the traditional group experienced a greater increase in scores from beginning to end of course. There was a significant interaction effect for the cognitive competence subscale where ratings increased by more than half a standard deviation for the traditional group but declined slightly for the online group. Given that face-to-face students gave higher ratings on the easiness subscale, it makes sense that their perceptions of their competence would also increase. For the remaining subscales of value, interest, and effort there were no significant interaction effects or main effects for course section, but there were significant main effects for time with ratings decreasing across all groups; the researchers note that this is consistent with previous research (e.g., Schau & Emmioğlu, 2012). Overall, the changes observed in SATS-36 subscale scores are in favor of the traditional students compared to the online students. However, they note that the SATS-36 was not previously validated for use with online students.

Suanpang et al. (2004) also compared changes in affect, cognitive competence, perceptions of value, and perceptions of easiness using the SATS-24 in online and face-to-face statistics students. Using a repeated measures analysis of variance, all time (beginning versus end of semester) by mode (online versus face-to-face) interactions were statistically significant. The online students' ratings tended to increase over time whereas face-to-face students' ratings tended to decrease slightly or remain unchanged. These results are different from those observed in Gundlach et al.'s (2015) study. The student populations that participated in the two studies were different, which may be impacting some of their results. In Gundlach et al.'s study, the participants were enrolled in a statistical literacy course offered by a university in the United States whereas in Suanpang et al.'s study, participants were enrolled in a business statistics course offered by a university in Thailand. Data concerning the ages of participants were not available in either study. In Section 2.2, we will present differences between traditional-aged students and adult learners; this may explain some of the variation in the results of these two studies as well.

In addition to attitudes towards statistics, the present study also examines statistics anxiety. Many students enter introductory statistics courses with feelings of anxiety (Onwuegbuzie, 2004; Zeidner, 1991) which may impact their course performance (Macher, Paechter, Papousek, & Ruggeri, 2012; Malik, 2015; Zare, Rastegar, & Hosseini, 2011). High anxiety may interfere with a student's ability to focus on the course content and to learn (Hanoch & Vitouch, 2004). Through interviews with introductory statistics students, Malik (2015) observed that students with high statistics anxiety were not confident in their abilities to succeed, which led to a lack of persistence. Anxiety has also been cited as a cause of procrastination in statistics courses (Malik, 2015; Onwuegbuzie, 2004). Again, the majority of this research has been conducted in face-to-face courses.

Macher et al. (2012) examined the relationships between anxiety and performance in an undergraduate statistics course. They measured statistics anxiety and trait anxiety (i.e., a "general anxiety proneness" (p. 484)) one week before a final exam. There was a negative relationship between statistics anxiety and final exam scores ( $r = -0.211$ ,  $p = 0.010$ ,  $N = 147$ ) and a positive relationship between statistics anxiety and procrastination ( $r = 0.261$ ,  $p = 0.001$ ,  $N = 147$ ). Statistics anxiety was most closely related to trait anxiety ( $r = 0.541$ ,  $p < 0.001$ ,  $N = 147$ ); according to the researchers, "Trait anxiety seems to foster the development of statistics anxiety, but the two concepts have a shared as well as an unshared component" (p. 492). Thus, although there is a strong relationship between trait anxiety and statistics anxiety, the two constructs are not identical.

The instrument used in the present study is an abbreviated form of the Statistical Anxiety Rating Scale (STARS) which consists of six subscales: Worth of Statistics, Self-Concept, Fear of Statistics Teachers (also known as attitudes towards statistics teachers), Interpretation Anxiety, Test Anxiety, and Asking for Help Anxiety. This instrument was selected because it addresses attitudes towards statistics in the first three subscales and statistics anxiety in the last three subscales (Cruise, Cash, and Bolton, 1985). The structure of the STARS was evaluated by DeVaney (2016) using a sample of online graduate students. His confirmatory factor analysis supported the use of the six-factor structure. Although researchers may not agree on all of the aspects of statistics attitudes to measure, they do tend to agree that it is a multidimensional construct (for a review of additional instruments that measure statistics attitudes, see Nolan, Beran, and Hecker, 2012).

In a study using a sample of students in an online undergraduate-level introductory statistics course, Zimmerman and Johnson (2017) developed and validated an abbreviated form of the STARS. They compared a one-factor, two-factor, and six-factor structure and concluded that a six-factor structure was most appropriate. In addition to examining the structure of an abbreviated form of the STARS, they examined differences between students who successfully complete the course and those who did not. Successful course completion was defined as finishing the course with a grade of D or higher. From a multivariate analysis of variance (MANOVA), there were not statistically significant differences between the STARS ratings at the beginning of the semester of students who did and did not successfully complete the course (Wilks' Lambda = 0.922,  $F(6, 316) = 0.624$ ,  $p = 0.711$ , partial eta squared = 0.012). Although there were no statistically significant differences, they did note that students who completed the course did tend to give more positive attitude ratings and lower anxiety ratings (Cohen's  $d$  ranging from 0.115 to 0.194). That study did not include any measure of achievement such as final exam score or overall course grade.

The lack of statistically significant differences in the STARS subscale scores of students who did and did not successfully complete the online undergraduate-level introductory statistics course in Zimmerman and Johnson's (2017) study was a bit surprising given that other studies had found connections between anxiety and success in introductory statistics courses (e.g., Macher et al., 2012; Malik, 2015; Zare et al., 2011). A major difference in the study by Zimmerman and Johnson (2017) and these other previous studies, is that Zimmerman and Johnson's sample was drawn from online sections of an introductory course. In the next section we discuss differences between students enrolled via the online campus and face-to-face campus.

## **2.2. FACE-TO-FACE AND ONLINE LEARNERS**

Compared to students enrolled in face-to-face courses, online students are more likely to be non-traditional, adult learners. Compared to traditional-aged students, adult learners tend to have more responsibilities outside of their coursework (Ko & Rossen, 2010). Adult learner status is defined by more than just age; characteristics such as parenthood, marital status, employment status, and military experience can also be used to classify individuals as adult learners (Hansman & Mott, 2010).

Students may choose to take online courses because they offer more flexibility than most face-to-face courses. Instead of regularly scheduled course meetings that require students to travel to a physical campus to attend in-person meetings, most online courses have weekly lessons with activities that can be completed asynchronously. Students may need to log on to the course multiple times a week, but there is great flexibility in terms of when during the week the student is present in the course. This can be appealing to adult learners (Conceição, 2007; Globokar, 2010). This added flexibility may make online courses appear to be easier than face-to-face courses, leading online students to overrate their abilities at the beginning of the semester (Dobbs, Waid-Lindberg, & del Carmen, 2017; Hoskins, 2014). In reality, online courses may be more demanding than face-to-face as they require more self-discipline to stay on schedule (Globokar, 2010; Wyatt, 2005).

Completion rates in online courses are often lower than for face-to-face courses, though the differences vary between courses and institutions. In an online introductory statistics course, McLaren (2004) compared online and face-to-face students' course completion rates. She found that the face-to-face students were more likely to complete the course (as opposed to dropping or "vanishing") compared to online students ( $\chi^2(2) = 51.701, p < 0.001$ ). Whereas overall completion rates may be lower in online courses, there may be interactions with other variables, such as adult learner status. Although traditional students may be more likely to succeed in face-to-face courses, adult learners may be more likely to succeed in online courses (Wladis, Conway, & Hachey, 2015).

Given that online learners tend to be adult learners with more going on in their lives beyond their schooling, it is hypothesized that attitudes and anxieties will be less powerful predictors of success in online students compared to face-to-face students. The purpose of this study was to examine how attitudes and anxiety can predict final exam grades and course completion in online and face-to-face sections of an undergraduate-level introductory statistics course. There were two primary research questions:

- (1) How can attitudes and anxieties be used to predict final exam scores, and does the relationship differ for online versus face-to-face introductory statistics students?
- (2) How can attitudes and anxieties be used to predict whether a student successfully completes the course, and does this relationship differ for online versus face-to-face introductory statistics students?

## **3. METHODS**

### **3.1. PARTICIPANTS AND CAMPUS INFORMATION**

Participants were 1,112 students enrolled in a four-credit undergraduate-level introductory statistics course with a lab component. This included 655 students from three face-to-face sections (all different instructors) and 457 students from 15 online sections (13 different instructors). Online course sections averaged 35 students per section, whereas face-to-face sections averaged around 78 students per lab section and more than 200 per lecture section. Approximately 525 online students were invited to

participate in the study and approximately 1,000 face-to-face students were invited to participate. This equates to an 87% participation rate from students in the online sections and a 65.5% participation rate from students in the face-to-face sections.

Demographics for the students involved in this study are not available, but in general, students attending classes on campus at this institution are primarily traditional students in the 18 to 22 year age range, whereas the majority of online students are adult learners. The average age of an undergraduate student from the online campus is 31 years. In terms of sex, 51.9% of students enrolled via the online campus are female and 46.7% enrolled via the face-to-face campus are female.

The online sections of the course were taught completely at a distance; all required activities were asynchronous, though there were optional live sessions with peer-tutors and some instructors offered to speak with students over the telephone or via video conferencing (e.g., Skype). The face-to-face sections of the course consisted of two lectures per week and two labs meetings per week. The labs were taught by graduate student teaching assistants. Both the online and face-to-face sections of the course used ANGEL as their course management system. The online and face-to-face sections of the course had the same stated course objectives and are treated as identical by the university; a student's transcript does not designate whether he or she completed the course online or face-to-face at one of the University's physical campuses.

### 3.2. INSTRUMENTATION

An abbreviated form of the Statistical Anxiety Rating Scale (STARS; Cruise et al., 1985; see also Hanna, Shevlin, and Dempster, 2008) was validated in a previous semester (Zimmerman & Johnson, 2017). The full version of the STARS consists of 51 items, which was judged to be too long to be completed by students in the present study. For the abbreviated scale, three items were selected on each of the six subscales: Test Anxiety, Asking Anxiety, Interpretation Anxiety, Worth of Statistics, Attitudes Toward Statistics Teachers, and Self-Concept, resulting in a total of 18 items. Anxiety items (Test Anxiety, Asking Anxiety, and Interpretation Anxiety Subscales) are measured using a five-point scale ranging from 1 (*no anxiety*) to 5 (*very strong anxiety*). Items on the remaining subscales (Worth Of Statistics, Attitudes Toward Statistics Teachers, and Self-Concept) are measured using a Likert rating scale from 1 (*strongly agree*) to 5 (*strongly disagree*). No items were reverse coded. The items on the abbreviated STARS are presented in Table 1.

*Table 1. Abbreviated STARS items*

Subscale	Item stem
Test Anxiety	Studying for an examination in a statistics course
	Doing an examination in a statistics course
	Waking up in the morning on the day of a statistics test
Asking Anxiety	Contacting my statistics instructor for help with material I am having difficulty understanding
	Asking one of my instructors for help in understanding a printout
	Asking a fellow student for help in understanding a printout
Interpretation Anxiety	Making an objective decision based on empirical data
	Reading a journal article that includes some statistical analyses
	Trying to understand the statistical analyses described in the abstract of a journal article
Worth of Statistics	I feel statistics is a waste
	I wish the statistics requirement would be removed from my academic major
	I am never going to use statistics
Attitudes Towards Statistics Teachers	Statistics teachers are so abstract they seem inhuman
	Statistics teachers communicate in a different language
	Statisticians are more number oriented than they are people oriented
Self-Concept	I cannot even understand high school math; I don't see how I can possibly do statistics
	Since I never enjoyed math, I do not see how I can enjoy statistics
	I do not have enough brains to get through statistics

### 3.3. PROCEDURES

During the first week of class, participants completed the abbreviated form of the STARS. For online students this was completed before they began the first graded lesson of the course. For face-to-face students this was during their first lab meeting; depending on the students' schedule this could be before or after the first lecture. After the end of the semester, final exam scores were recorded and whether or not each student successfully completed the course with a grade of D or higher was recorded. This was the lowest possible passing grade. A student who did not successfully complete a course either received a failing grade or dropped the course. These two variables (final exam scores and successful course completion) are used as measures of student success.

Final exam scores are used as measure of success for three reasons: (1) All final exams were designed to assess proficiency of all learning objectives; (2) Assignment categories and weights have slight variations across sections; and (3) Final course grades were not available for all students. Students in the online sections of the course all took the same final exam. The face-to-face sections of the course had some shared questions and some questions that were unique to each section. Thus, there were four versions of the exam: one for each of the three face-to-face sections and one for all online sections. All final exams were comprised entirely of multiple-choice questions and all sections of the course had the same stated learning objectives.

### 3.4. ANALYSIS METHODS

After verifying the appropriateness of the abbreviated STARS measurement instrument (Section 4.1), and considering descriptive statistics of the STARS subscales (Section 4.2), we will address the two research questions. In Section 4.3, the relationship of students' attitudes and anxieties to final exam scores is assessed using pairwise correlations and linear regression analysis. In Section 4.4, chi-square tests are used to detect differences in the completion rate between online and face-to-face students. This is followed by an analysis of the relationship between a student's attitudes and anxieties and whether or not the student completes the course. Both logistic regression and random forest models are applied to the data; although logistic regression has good theoretical properties and easy interpretation, it does not work well in all situations. For example, unlike logistic regression models, random forests are non-parametric, provide greater flexibility in defining the relationship between the predictors and the response, and do not overfit data. Furthermore, in the case of imbalanced classes, logistic regression often will never or rarely predict that an observation will fall into the smaller class (in this case, the smaller class is the class of students who did not complete the class, compared to those who did complete). In situations like these, random forests may outperform logistic regression in prediction but at the cost of interpretability; however, random forests do provide metrics to identify the most important terms in the model, one of which is discussed in Section 4.4.

A random forest is an ensemble method that applies ideas of bootstrap aggregating ("bagging") and random sampling of predictors to classification and regression trees. In bagging, individual decision trees are fit using bootstrapped samples and the predicted class for a given observation is the majority class from all the trees; bagging is used to reduce the overall variance of the model. A random forest takes bagging one step further by using only a random subset of predictors as candidates for splitting at each node when building the trees; this leads to decorrelation of trees, subsequently reducing variance further.

For more on random forest and other tree-based methods, see James et al. (2013) and Breiman (2001). For a comparison of logistic regression and random forest in the case of unbalanced class sizes, see Muchlinski et al. (2015).

## 4. RESULTS

Before abbreviated STARS subscale scores could be compared between the online and face-to-face groups, we examine the structure of the scale using measurement invariance techniques. Following this analysis, we computed descriptive statistics for the two groups separately, then addressed the two research questions.

#### 4.1. MEASUREMENT INVARIANCE

Before any descriptive or inferential statistics are performed using subscale scores, it is necessary to examine the factor structure of the data collected from these two groups. The abbreviated form of the STARS was designed to be consistent with the six-factor structure of the full form. We conducted confirmatory factor and measurement invariance analyses to compare the six-factor structure of the abbreviated STARS for online and face-to-face students; these results were statistically significant ( $\chi^2(12) = 23.814, p = 0.0216$ ). The factor loadings were very similar and we determined that it was appropriate to compare the subscale scores of the two groups. Model-fit results for the six-factor model with the groups combined were statistically significant ( $\chi^2(120) = 616.6, p < 0.001, RMSEA = 0.061, CFI = 0.947$ ), however this may be due to high power given the large sample size. RMSEA and CFI values suggest that the model fit is sufficient (Byrne, 2009).

#### 4.2. STARS DESCRIPTIVE STATISTICS

Descriptive statistics concerning the six abbreviated STARS subscales for the online and face-to-face groups from the first week of class are given in Table 2. Confidence intervals for the difference in means were calculated (using unpooled variances) to compare each subscale score of the online and face-to-face groups. A confidence level of 0.99 was used to account for multiple estimation; this is the approximate confidence level for a significance level of 0.05 with a Bonferroni adjustment. Statistically significant differences were found between the two groups in terms of ratings on the Asking for Help Anxiety and Worth of Statistics subscales. Students in the online sections of the course had lower anxiety for Asking for Help and gave higher ratings on the Worth of Statistics subscale, but in all cases the expected difference is no more than 0.4 points.

Table 2. Descriptive statistics for STARS subscales by campus

	Online			Face-to-Face			99% CI (Online-FTF)
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	
Test Anxiety	447	3.154	1.019	650	3.104	0.901	(-0.105, 0.204)
Asking for Help Anxiety	452	1.923	0.915	651	2.134	0.854	(-0.353, -0.071)*
Interpretation Anxiety	442	2.137	0.836	645	2.130	0.793	(-0.124, 0.137)
Worth of Statistics	440	3.892	0.840	643	3.729	0.744	(0.035, 0.292)*
Attitudes Toward Statistics Teachers	447	3.711	0.731	643	3.640	0.708	(-0.046, 0.185)
Self-Concept	453	3.896	0.995	649	3.840	0.964	(-0.099, 0.211)

\*  $p$ -value < 0.05

Tolerance and variance inflation factor (VIF) values for the six abbreviated STARS subscales are given in Table 3. Online and face-to-face students were combined for this analysis. Given the tolerance and VIF values, it was concluded that the overlap between the subscales was minimal enough that each of the six subscales may be included in the analyses for both research questions.

Table 3. Variance inflation among abbreviated STARS subscales

Subscale	Tolerance	VIF
Test Anxiety	0.522	1.915
Asking for Help Anxiety	0.732	1.365
Interpretation Anxiety	0.559	1.788
Worth of Statistics	0.543	1.843
Attitudes Toward Statistics Teachers	0.632	1.581
Self-Concept	0.512	1.954

#### 4.3. RESEARCH QUESTION #1

The first research question was “How can attitudes and anxieties be used to predict final exam scores, and does this relationship differ for online versus face-to-face introductory statistics students?”

There were four different final exams and the distribution of scores differed (Figure 1), so the final exam scores were standardized (centered and scaled) prior to analysis. Because all online sections took the same exam, and because evidence indicates the scores from the online sections follow approximately the same distribution (Kruskal-Wallis rank sum test,  $\chi^2 = 16.573$ ,  $p = 0.1664$ ), scores were standardized within exam version rather than section.

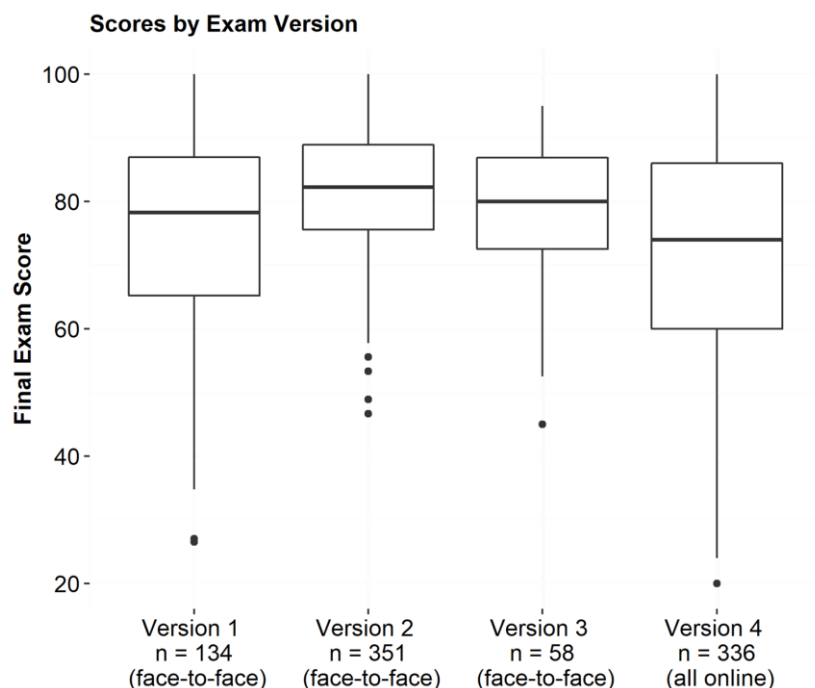


Figure 1. Final exam score by exam version

First we explored the pairwise correlations between the standardized scores and each of the subscales. These are summarized for each of the two groups in Table 4, along with the adjusted  $p$ -values using the Benjamini and Hochberg method. There were 336 online students and 543 face-to-face students who took the final exams. The last column provides the bootstrapped confidence interval using 10,000 bootstrap samples for the pairwise difference in correlations using a 0.99 confidence level. This higher confidence level was chosen as a conservative estimate to adjust for multiple comparisons as well as to offset the “overly confident” nature of bootstrap estimates.

Table 4. Pairwise correlation exam scores and STARS subscales by campus

Subscale	Online		Face-to-Face		99% bootstrap CI (Online-FTF)
	$r$	$p^a$	$r$	$p^a$	
Test Anxiety	-0.170	0.004	-0.146	0.009	(-0.193, 0.144)
Asking For Help	-0.068	0.260	-0.060	0.146	(-0.190, 0.182)
Interpretation Anxiety	-0.124	0.035	-0.098	0.026	(-0.214, 0.165)
Worth Of Statistics	0.203	0.0005	0.238	< 0.0001	(-0.203, 0.132)
Attitudes Towards Teachers	-0.010	0.854	0.194	< 0.0001	(-0.382, -0.023)*
Self-Concept	0.210	0.0005	0.298	< 0.0001	(-0.258, 0.083)

<sup>a</sup>  $p$ -values are BH-adjusted

\* Difference between online and face-to-face correlations significant at  $\alpha = 0.01$

As seen in Table 4, Test Anxiety, Interpretation Anxiety, Worth of Statistics, and Self-Concept are all correlated with standardized final exam scores for each group, with the latter two having the strongest correlations. However, some of these correlations, such as that between Interpretation Anxiety and final exam scores in the face-to-face group (-0.098), are too small to likely be of practical importance. The only practically significant difference between online and face-to-face students in the



correlation between standardized final exam scores and the STARS subscales is with Attitudes Towards Statistics Teachers; there is a stronger correlation among face-to-face students ( $r = 0.194$ ) and virtually no correlation among online students ( $r = -0.010$ ).

Pairwise correlations do not account for the relationship the scores have with the other subscales, so to bolster the analysis, linear regression was used to model the relationship between the dependent variable (final exam scores) and the independent variables, which included the campus (1 = online, 0 = face-to-face), STARS attitude and anxiety subscale ratings, and the interactions between the campus and each subscale.

Stratified random sampling (by campus, non-missing test scores) was used to divide the data into two equally-sized sets for model selection ( $n = 440$ ) and inference ( $n = 439$ ). Using the model selection dataset, the full model was found to be significant ( $F = 5.598, p < 0.0001$ ). Furthermore, a general linear  $F$ -test testing with a reduced model including only main effect terms showed that there were no significant interaction terms ( $F = 1.247, p = 0.281$ ). Thus, final model selected includes only the main effect terms for campus and the six STARS subscale items ( $F = 6.422, p < 0.0001$ ). This reduced model was fit to the inference dataset to estimate parameters and test for significance; the results are provided in Table 5. A residual analysis suggests that no model assumptions were violated, but the predictive power is low ( $R^2 = 0.09445$ ).

Table 5. Linear regression of standardized final exam scores on inference data

Term	Coeff.	Std. Error	$t$	$p$ -value
Campus	0.122	0.092	1.320	0.1874
Test Anxiety	-0.078	0.066	-1.184	0.2370
Asking for Help Anxiety	0.017	0.059	0.288	0.7738
Interpretation Anxiety	-0.021	0.074	-0.291	0.7715
Worth of Statistics	0.183	0.076	2.414	0.0162*
Attitudes Towards Statistics Teachers	-0.018	0.080	-0.222	0.8242
Self-Concept	0.136	0.066	2.055	0.0405*

\* $p < 0.05$

As seen in Table 5, Worth of Statistics and Self-Concept are each statistically significant predictors at the  $\alpha = 0.05$  level. Each contribute positively to exam scores: All else equal, as a student places more value in statistics or has higher confidence in his or her ability to do well in a statistics course, his or her expected final exam score increases.

The lack of any significant interaction terms, however, suggests that the relationship between the STARS subscales and the final exam scores do not differ between online and face-to-face students. This is in contrast to what was found in the pairwise correlation analysis; this suggests that the difference seen there can be explained by the contribution of the other subscales to the final exam scores.

#### 4.4. RESEARCH QUESTION #2

The second research question was “How can attitudes and anxieties be used to predict whether a student successfully completes the course, and does this relationship differ for online versus face-to-face introductory statistics students?” Tabulation for successful course completion and campus is in Table 6. The overall completion rate in the dataset is 83.7%; the rate is higher for face-to-face students compared to online students, a difference which was found to be statistically significant ( $\chi^2 = 8.94, p = 0.0028, 95\% \text{ CI for difference} = (2.2\%, 11.6\%)$ ). However, due to the difference in audience for online compared to face-to-face courses at this institution, this is not surprising. There was no significant difference in the completion rate across the online sections ( $\chi^2 = 5.81, p = 0.9254$ ) but there was a moderately significant difference in the completion rate across the face-to-face sections ( $\chi^2 = 7.19, p = 0.027$ ), suggesting there may be an instructor or specific lesson-plan effect for face-to-face courses.

A logistic regression model was fit to estimate the probability of successful completion based on campus (1 = online, 0 = face-to-face), STARS attitude and anxiety subscale ratings, and the interactions between the campus and each subscale. Although the full model is significant ( $D = 24.5, p = 0.0268$ ),

Table 6. Successful completion by campus counts and within-campus percentages

	Complete		Incomplete		Total Count
	Count	Within-campus %	Count	Within-campus %	
Online	364	79.6	93	20.4	457
Face-to-face	567	86.6	88	13.4	655
All	931	83.7	181	16.3	1112

none of the terms except campus were statistically significant; testing against a reduced model with only campus as a predictor showed that neither the subscales nor the interaction of the subscales with campus were significant ( $D = 16.052$ ,  $p = 0.189$ ). However, under both the full model and reduced model, all of the participants in the present study were classified into the category of “successful course completion,” making it a poor model for prediction, possibly due to the imbalanced class sizes or the complex nature of the relationship between the predictors and the response.

In an attempt to improve prediction, a random forest model with 10,000 trees was used to predict successful completion from campus and each of the six STARS subscales. “Out-of-bag” (OOB) samples—that is, those data points that were not used in building each of the trees—can be used to estimate the prediction accuracy of each tree. The average out-of-bag classification error rate for the trees was 17%, but the error rate for the forest when applied to the original data set is just 1.6%, with all “successful completion” participants correctly classified and 90.4% of “incomplete” participants correctly classified.

The mean decrease in accuracy (MDA) values for each predictor are given in Figure 2. The MDA of a predictor variable is a scaled metric for the decrease in the accuracy (or equivalently, increase in error) of the OOB samples that would occur if the variable was permuted. Specifically, for each tree, the difference in the OOB error rate before and after permuting the variable is averaged over all trees and then normalized by the standard deviation of the differences. It is not the measure of the difference in the overall forest accuracy but does provide a measure of importance of the variable in creating the individual trees, thus providing a natural way to rank predictors for further investigation.

Interestingly, campus contributed the least to accuracy in tree predictions by the MDA metric, whereas Test Anxiety and Self-Concept had the highest values, suggesting they are more powerful predictors than the others in this model.

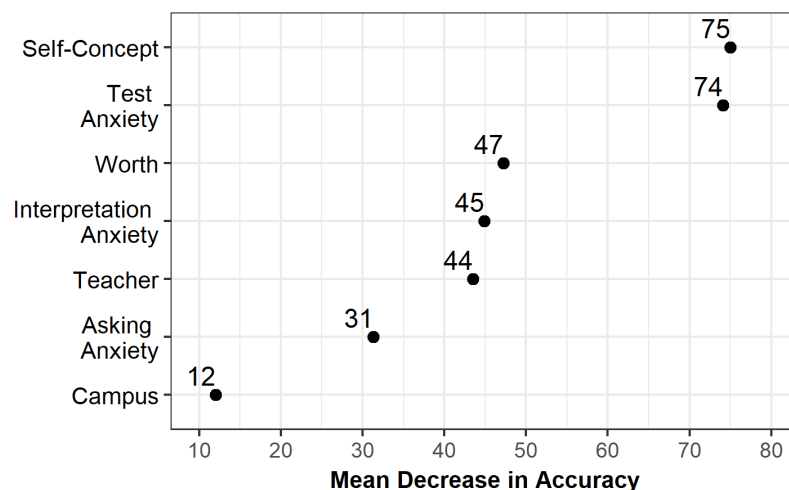


Figure 2. Mean decrease in accuracy for the random forest model

Random forests were also grown separately for the online and face-to-face groups. Using the same MDA metric, the most important predictor for classification in the online format was Test Anxiety followed by Self-Concept and Interpretation Anxiety, whereas in the face-to-face group, the leading predictor was Self-Concept followed by Test Anxiety. This suggests that there may be a difference

between online and face-to-face introductory statistics students in how these subscales contribute to whether a student successfully completes the course, but further research is necessary.

## 5. DISCUSSION

Comparing the mean subscale scores for the two groups, online students had lower levels of anxiety for asking questions and more positive perspectives on the worth of statistics compared to face-to-face students. These differences may be due to demographic differences between the two groups (e.g., age), though those data were not available to analyze. However, in general, online students are more likely to be adult learners whereas students in the face-to-face course are primarily traditional students. The lower levels of anxiety for asking questions may also be present due to the anonymity associated with being online or because students are afforded the option to spend more time crafting questions before posting them to an online discussion board or sending them via email. The fact that the online sections of the course were smaller than the face-to-face sections may also produce a small-class atmosphere that reduces anxiety of asking questions. The more positive perceptions of the worth of statistics are likely due to differences in the populations: Adult learners have more life experiences and therefore may see more practical applications of statistics, whereas traditional students may not yet realize all of the value of understanding statistics.

Linear regression analysis shows that there exists a relationship between the Worth of Statistics and Self-Concept abbreviated STARS subscale scores and standardized final exam scores, but that the predictive power of the overall model is low. The lack of significant interaction terms suggests that this relationship is similar for face-to-face and online students, contradicting the original hypothesis. Although pairwise correlation tests did suggest that there may be a difference in the correlation of final exam scores and Attitudes Towards Statistics Teachers in face-to-face versus online students, this did not account for the variation in scores explained by the other subscales. However, this finding does suggest that further research may be warranted.

Given the design of this study we do not know whether this difference, if real, is due to differences in the students who tend to take online versus face-to-face courses (i.e., adult learners versus traditional students) or whether it is related to differences in the mode of instruction. As originally hypothesized, there may be more lurking variables influencing online learners because they tend to be adults with many additional responsibilities that can impact their course performance. Considering the impact of the mode of instruction, it is possible that face-to-face students' preconceptions about statistics instructors in general may impact their learning due to the fact that the content is being delivered through the instructor in lectures. In other words, in face-to-face courses learning may be negatively impacted if students enter the course with bias against statistics teachers. In online classes the content is more often delivered through online notes and asynchronous communication with the instructor (e.g., prerecorded videos, discussion boards) which may explain why there is not a statistically significant correlation between final exam scores and attitudes towards teachers in the online students. Additional research is needed to determine whether this difference is real or, as suggested by the linear regression, is spurious, and where the causation lies if it is real. Although it is hypothesized that observed differences are due to differences in the audience, over which instructors may have relatively little control, if there exist other controllable explanations, those could be manipulated by instructors to create more optimal learning environments.

As expected, there is a significant difference in the course completion rate by campus, with online students being less likely to successfully complete a course with a D or higher (consistent with McLaren, 2004). In the logistic regression model, none of the subscales contributed significantly to the model and the model had poor predictive power. Increased prediction accuracy was obtained using a random forest model, with Test Anxiety and Worth of Statistics contributing the most to predictive accuracy. Interestingly, if random forests were fit to online and face-to-face participants separately, Test Anxiety was the major contributor for online group whereas Self-Concept was the leading predictor for the face-to-face group. This difference warrants further study as it suggests that the predictive contributions of the subscale scores to course success may differ for online versus face-to-face students.

The findings that Self-Concept and Worth of Statistics show up as factors contributing to student success as measured by final exam scores and completion rate are aligned with the expectancy-value

model of achievement which describes motivation in terms of the subjective value of the task and expectations of success (Wigfield & Eccles, 2000). Future research will apply this theory in an attempt to better explain student success.

In the present study the STARS was only administered at the beginning of the course. In the future, the survey should also be given at the end of the course. This will allow for a comparison of online and face-to-face students over time. Additionally, future research should collect demographic data from participants including age and experience with online learning.

## 6. CONCLUSIONS

Previous research conducted in face-to-face statistics courses had observed relations between attitudes towards statistics and achievement (Emmioglu & Capa-Aydin, 2012) as well as statistics anxiety and achievement (Macher et al., 2012; Malik, 2015). Those relations were not observed in the present study in terms of final exam scores, but it was observed that test anxiety may be related to successful course completion by online students. Because the exams differed for face-to-face and online sections of this course it is difficult to determine whether the relationship, if real, is due to course/exam delivery, student characteristics, or differences in the exams themselves.

The present study also found that self-concept and students' perceived worth of statistics can be used to predict final exam scores but that this relationship was not significantly different by campus (online versus face-to-face). However, the random forest models suggest that there may be a difference by campus in which STARS subscales best predict whether a student successfully completes the course with a D or better.

Initial findings suggest that there does not appear to be significant differences in the relationship between the final exam scores and student attitudes and anxieties for online and face-to-face groups, but there may be a difference when it comes to course completion. Thus, some of the findings do provide evidence to support the hypothesis that attitudes and anxieties may be less powerful predictors of success in online students compared to face-to-face students, possibly due to more lurking variables impacting the success of online students. The differences in the course curricula, including the exams and the weighting of assignments, provide inferential challenges so more research is needed to better examine these relationships.

## ACKNOWLEDGEMENTS

The authors would like to thank all of the instructors of the online introductory statistics course that was featured in this study for their participation. They would also like to thank the anonymous reviewers of this paper, Glenn Johnson, and Durland Shumway for their valuable feedback and support.

## REFERENCES

- Allen, I. E., & Seaman, J. (2016). *Online report card: Tracking online education in the United States*. Babson Survey Research Group and Quahog Research Group, LLC.  
[Online: <http://onlinelearningsurvey.com/reports/online-report-card.pdf> ]
- Allen, I. E., & Seaman, J. (2017). *Digital learning compass: Distance education enrollment report 2017*. Babson Survey Research Group, e-Literate, and WCET.  
[Online: <http://onlinelearningsurvey.com/reports/digital-learning-compass-enrollment-2017.pdf> ]
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.  
[Online: <https://doi.org/10.1023/A:1010933404324> ]
- Byrne, B. M. (2009). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York: Psychology Press, Taylor & Francis Group.
- Cochran, J. D., Campbell, S. M., Baker, H. M., & Leeds, E. M. (2014). The role of student characteristics in predicting retention in online courses. *Research in Higher Education*, 55(1), 27–48.  
[Online: <http://doi.org/10.1007/s11162-013-9305-8> ]

- Conceição, S. C. O. (2007). Understanding the environment for online teaching. In S. C. O. Conceição (Ed.), *Teaching strategies in the online environment* (pp. 5–11). Hoboken, NJ: Wiley Periodicals, Inc.
- Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology*, *57*, 345–374.
- Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. *Proceedings of the American Statistical Association, Section on Statistical Education* (pp. 92–97). Alexandria, VA: American Statistical Association.
- DeVaney, T. A. (2016). Confirmatory factor analysis of the Statistical Anxiety Rating Scale with online graduate students. *Psychological Reports*, *118*(2), 565–586.
- Dobbs, R. R., Waid-Lindberg, C. A., & del Carmen, A. (2017). Students' perceptions of online courses: The effect of online course experience. *Quarterly Review of Distance Education*, *18*(1), 93–109.
- Emmioglu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, *11*(2), 95–102.  
[Online: [https://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Emmioglu.pdf](https://iase-web.org/documents/SERJ/SERJ11(2)_Emmioglu.pdf) ]
- Gaytan, J. (2013). Factors affecting student retention in online courses: Overcoming this critical problem. *Career and Technical Education Research*, *38*(2), 145–144.  
[Online: <http://doi.org/10.5328/cter38.2.147> ]
- Globokar, J. L. (2010). *Introduction to online learning: A guide for students*. Los Angeles: Sage Publications.
- Gundlach, E., Richards, K. A. R., Nelson, D., & Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class. *Journal of Statistics Education*, *23*(1).  
[Online: <http://ww2.amstat.org/publications/jse/v23n1/gundlach.pdf> ]
- Hanna, D., Shevlin, M., & Dempster, M. (2008). The structure of the Statistics Anxiety Rating Scale: A confirmatory factor analysis using UK psychology students. *Personality and Individual Differences*, *45*(1), 68–74.
- Hanoch, Y., & Vitouch, O. (2004). When less is more: Information, emotional arousal and the ecological reframing of the Yerkes-Dodson Law. *Theory & Psychology*, *14*(4), 427–452.
- Hansman, C. A., & Mott, V. W. (2010). Adult learners. In C. E. Kasworm, A. D. Rose, & J. M. Ross-Gordon (Eds.), *Handbook of adult and continuing education* (2010 ed., pp. 13–23). Thousand Oaks, CA: SAGE Publications, Inc.
- Hoskins, B. J. (2014). The anatomy of online offerings. *The Journal of Continuing Higher Education*, *62*(1), 62–64.  
[Online: <https://doi.org/10.1080/07377363.2014.877767> ]
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). Tree-based methods. In *An introduction to statistical learning with applications in R*, (pp. 302–324). New York: Springer
- Ko, S., & Rossen, S. (2010). *Teaching online: A practical guide*. New York: Routledge.
- Macher, D., Paechter, M., Papousek, I., & Ruggeri, K. (2012). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European Journal of Psychology of Education*, *27*(4), 483–498.  
[Online: <http://doi.org/10.1007/s10212-011-0090-5> ]
- Malik, S. (2015). Undergraduates' statistics anxiety: A phenomenological study. *The Qualitative Report*, *20*(2), 120–133.
- McLaren, C. H., (2004). A comparison of student persistence and performance in online and classroom business statistics experiences. *Decision Sciences Journal of Innovative Education*, *2*(1), 1–10.
- Muchlinski, D., Siroky, D., He, J., & Kocher, M. (2015). Comparing random forest with logistic regression for predicting class-imbalanced Civil War onset data. *Political Analysis*, 1–17.  
doi:10.1093/pan/mpv024
- Nolan, M. M., Beran, T., & Hecker, T. G. (2012). Surveys assessing students' attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, *11*(2), 103–123.  
[Online: [https://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Nolan.pdf](https://iase-web.org/documents/SERJ/SERJ11(2)_Nolan.pdf) ]
- Onwuegbuzie, A. J. (2004). Academic procrastination and statistics anxiety. *Assessment and Evaluation in Higher Education*, *29*(1), 3–19.

- Pearl, D. K., Garfield, J. B., delMas, R., Groth, R. E., Kaplan, J. J., McGowen, H., & Lee, H. S. (2012). *Connecting research to practice in a culture of assessment for introductory college-level statistics*. [Online: [https://www.causeweb.org/cause/archive/research/guidelines/ResearchReport\\_2012.pdf](https://www.causeweb.org/cause/archive/research/guidelines/ResearchReport_2012.pdf) ]
- Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, 11(2), 57–71.  
[Online: [https://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Ramirez.pdf](https://iase-web.org/documents/SERJ/SERJ11(2)_Ramirez.pdf) ]
- Suanpang, P., Petocz, P., & Kalceff, W. (2004). Student attitudes to learning business statistics: Comparison of online and traditional methods. *Journal of Educational Technology & Society*, 7(3), 9–20.
- Schau, C. (2003). *Survey of Attitudes Towards Statistics (SATS-36)*.  
[Online: <http://www.evaluationandstatistics.com> ]
- Schau, C., & Emmioğlu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? *Statistics Education Research Journal*, 11(2), 86–94.  
[Online: [https://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Schau.pdf](https://iase-web.org/documents/SERJ/SERJ11(2)_Schau.pdf) ]
- Schau, C., Millar, M., & Petocz, P. (2012). Research on attitudes towards statistics. *Statistics Education Research Journal*, 11(2), 2–5.  
[Online: [https://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Editorial.pdf](https://iase-web.org/documents/SERJ/SERJ11(2)_Editorial.pdf) ]
- Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, 55(5), 868–875.
- Vanhoof, S., Kuppens, S., Castro Sotos, A. E., Verschaffel, L., & Onghena, P. (2011). Measuring statistics attitudes: Structure of the Survey of Attitudes Toward Statistics (SATS-36). *Statistics Education Research Journal*, 10(1), 25–51.  
[Online: [https://iase-web.org/documents/SERJ/SERJ10\(1\)\\_Vanhoof.pdf](https://iase-web.org/documents/SERJ/SERJ10(1)_Vanhoof.pdf) ]
- Vogt, W. P. (2005). Attitude. In *Dictionary of statistics & methodology: A nontechnical guide for the social sciences* (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 69–81.  
[Online: <http://doi.org/10.1006/ceps.1999.1015> ]
- Williams, A. S. (2013). Worry, intolerance of uncertainty, and statistics anxiety. *Statistics Education Research Journal*, 12(1), 48–59.  
[Online: [https://iase-web.org/documents/SERJ/SERJ12\(1\)\\_Williams.pdf](https://iase-web.org/documents/SERJ/SERJ12(1)_Williams.pdf) ]
- Williams, A. S. (2015). Statistics anxiety and worry: The roles of worry beliefs, negative problem orientation, and cognitive avoidance. *Statistics Education Research Journal*, 14(2), 53–75.  
[Online: [https://iase-web.org/documents/SERJ/SERJ14\(2\)\\_Williams.pdf](https://iase-web.org/documents/SERJ/SERJ14(2)_Williams.pdf) ]
- Wladis, C., Conway, K. M., & Hachey, A. C. (2015). The online STEM classroom - who succeeds? An exploration of the impact of ethnicity, gender, and non-traditional student characteristics in the community college context. *Community College Review*, 43(2), 142–164.  
[Online: <http://doi.org/10.1177/0091552115571729> ]
- Wyatt, G. (2005). Satisfaction, academic rigor and interaction: Perceptions of online instruction. *Education*, 125(3), 460–468.
- Zare, H., Rastegar, A., & Hosseini, S. M. D. (2011). The relation among achievement goals and academic achievement in statistics: The mediating role of statistics anxiety and statistics self-efficacy. *Procedia, Social and Behavioral Sciences*, 30, 1166–1172.
- Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, 61(3), 319–328.
- Zimmerman, W. A., & Johnson, G. (2017). Exploring factors related to completion of an online undergraduate-level introductory statistics course. *Online Learning Journal*, 21(3), 191–205.

WHITNEY ALICIA ZIMMERMAN  
The Pennsylvania State University  
Department of Statistics  
325 Thomas Building  
University Park, PA16083  
USA