HOW ATTITUDES TOWARDS STATISTICS COURSES AND THE FIELD OF STATISTICS PREDICTS STATISTICS ANXIETY AMONG UNDERGRADUATE SOCIAL SCIENCE MAJORS: A VALIDATION OF THE STATISTICAL ANXIETY SCALE

Monique J. O'Bryant

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2017

APPROVED:

Prathiba Natesan, Major Professor
Mei Chang, Committee Member
Smita Mehta, Committee Member
Anthony Onwuegbuzie, Committee Member
Karthigeyan Subramanian, Committee Member
Abbas Tashakkori, Chair of the Department of Educational Psychology
Randy Bomer, Dean of the College of Education
Victor R. Prybutok, Dean of the Toulouse Graduate School

O'Bryant, Monique J. *How Attitudes towards Statistics Courses and the Field of Statistics Predicts Statistics Anxiety among Undergraduate Social Science Majors: A Validation of the Statistical Anxiety Scale*. Doctor of Philosophy (Educational Research), August 2017, 62 pp., 7 tables, 2 figures, references, 89 titles.

The aim of this study was to validate an instrument that can be used by instructors or social scientist who are interested in evaluating statistics anxiety. The psychometric properties of the English version of the Statistical Anxiety Scale (SAS) was examined through a confirmatory factor analysis of scores from a sample of 323 undergraduate social science majors enrolled in colleges and universities in the United States. In previous studies, the psychometric properties of the Spanish and Italian versions of the SAS were validated; however, the English version of the SAS had never been assessed. Inconsistent with previous studies, scores on the English version of the SAS did not produce psychometrically acceptable values of validity. However, the results of this study suggested the potential value of a revised two-factor model SAS to measure statistics anxiety. Additionally, the Attitudes Towards Statistics (ATS) scale was used to examine the convergent and discriminant validities of the two-factor SAS. As expected, the correlation between the two factors of the SAS and the two factors of the ATS uncovered a moderately negative correlation between examination anxiety and attitudes towards the course. Additionally, the results of a structural regression model of attitudes towards statistics as a predictor of statistics anxiety suggested that attitudes towards the course and attitudes towards the field of statistics moderately predicts examination anxiety with attitudes towards the course having the greatest influence. It is recommended that future studies examine the relationship between attitudes towards statistics, statistics anxiety, and other variables such as academic achievement and instructional style.

ACKNOWLEDGEMENTS

I would like to thank my major professor and committee chair, Prathiba Natesan, for her guidance, support, and encouragement throughout the doctoral program. It has been a long journey, and there is no way I would have ever made it this far without you. To Mei Chang, thank you for your support and offering me ideas to improve my dissertation and build on my research. To Smita Mehta, thank you for pushing me to think creatively in my dissertation sampling methods and all the program evaluation knowledge you've shared with me in my coursework. To Karthigeyan Subramanian, thank you for your encouragement and participation on this committee; I've learned so much from you in the area of qualitative methods and look forward to applying this skills in future research. To Anthony Onwuegbuzie, it has been an honor to have you on my committee and to work with such a profound researcher; thank you for all your support and involvement. To the Department of Educational Psychology, thank you for sharing your knowledge and training in my coursework. These skills have afforded me the opportunity to work in a variety of settings and will continue to lead me in my future endeavors. To my many classmates, friends, and co-workers that have supported me along the way, thank you for being a shoulder to cry on and for all the thoughtful words of encouragement. To my family, thank you for being understanding and supporting over the years. There is no way I could have ever reached this huge milestone without you. Thank you to my soon-to-be husband for being patient with me over the years and understanding when my time was committed to my studies. You believed in me and pushed me, even when I did not believe in myself. I am forever grateful for you. Lastly, appreciation is extended to Andreu Vigil-Colet for permission to use the Statistical Anxiety Scale and to Steven Wise for permission to use the Attitudes Toward Statistics Scale in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

HOW ATTITUDES TOWARDS STATISTICS COURSES AND THE FIELD OF STATISTICS PREDICTS STATISTICS ANXIETY AMONG UNDERGRADUATE SOCIAL SCIENCE MAJORS A VALIDATION OF THE STATISTICAL ANXIETY SCALE

Introduction

As part of the academic training in post-secondary schools, students in the social sciences often are required to take at least one statistics course regardless of their background in statistics or mathematics (Pan & Tang, 2004). These courses are essential because they might be the only formal statistical training that students receive (Onwuegbuzie, 2000). Furthermore, these courses provide the statistical background needed to conduct quantitative research studies and to comprehend empirical research (Birenbaum & Eylath, 1994). For students in non-mathematical disciplines, statistics courses often are associated with negative experiences and high levels of anxiety. Also, they are considered the most difficult courses in their programs (Chew & Dillon, 2014; Cruise, Cash, & Bolton, 1985; Onwuegbuzie, Leech, Murtonen, & Tahtinen, 2010). There are many aspects of a statistics course that might cause students to experience anxiety (Hanna, Shevlin, & Dempster, 2008). Statistics anxiety has been found to cause students to display feelings of depression, frustration, anger, apprehension, nervousness, worry, panic, stress, and emotionality (Onwuegbuzie, Da Ros, & Ryan, 1997). Likewise, statistics anxiety has been associated with the anxiety experienced by graduate students when writing their research proposals (Onwuegbuzie, 1997). Moreover, statistics anxiety has been shown to hinder performance in statistics courses for approximately 80% of graduate students (Onwuegbuzie, 2004).

Statistics anxiety was defined by Cruise et al. (1985) as “feelings of anxiety encountered when taking a statistics course or doing statistical analysis; that is, gathering, processing and interpreting data” (p. 92). Zeidner (1991) further depicted statistics anxiety as a form of performance anxiety represented by disturbing thoughts of worry, tension, and mental disorganization that arise when people are exposed to statistics content, instructional situations, or evaluative contexts. Onwuegbuzie, Da Ros, and Ryan (1997) later described statistics anxiety as an anxiety that occurs as a consequence of encountering statistics in any form or at any level. In addition, statistics anxiety has been defined as being situation-specific because the symptoms may only appear in a particular setting and time (Baloglu, 2003; Onwuegbuzie, 1999). Unfortunately, these definitions do not discuss the relationship statistics anxiety has with mathematics anxiety and with attitudes towards statistics (Chew & Dillon, 2014).

Statistics involves more than the manipulation of numbers and solving problems; it encompasses collecting and using data as well as applying the results from the data analysis to inform decisions (Cruise et al., 1985). Baloglu (2002) confirmed that statistics anxiety and mathematics anxiety were highly correlated ($r = .67$) but distinct constructs (only 45% shared variance). Thus, it is suggested by Chew and Dillon (2014) that statistics anxiety be redefined as a negative emotional state stimulated from any form of interaction with statistics and exacerbated by negative attitudes towards it. This negative feeling is associated with, but separate from, mathematics anxiety (Chew & Dillon, 2014).

When a person encounters a feeling of anxiety, the specific cause of their overall emotions might be difficult to identify (Williams, 2013). Because of this difficulty, over the years, several measures of statistics anxiety have been developed. One of the most widely used measures of statistics anxiety is the Statistical Anxiety Rating Scale (STARS; Cruise and Wilkins, 1980). This 51-item assessment was developed by Cruise and Wilkins (1980) to measure statistics anxiety via the following six subscales: (a) computational self-concept, (b) fear of asking for help, (c) fear of statistics teachers, (d) interpretation anxiety, (e) test and class anxiety, and (f) worth of statistics (Cruise & Wilkins, 1980). According to the STARS developers, computational self-concept refers to the level of anxiety that a student feels when solving mathematical problems as well as their perceptions of their ability to understand statistics. Fear of asking for help is concerned with the amount of anxiety felt when requesting help pertaining to statistics from another student or a teacher. Fear of the statistics teacher involves the student's perception of the statistics teacher. Interpretation anxiety is the anxiety that students encounter when they have to make a decision or to interpret statistics data. Test and class anxiety is the level of anxiety felt when students are involved in the statistics class or assessment. Finally, worth of statistics refers to the student's perception of the usefulness of statistics (Cruise et al., 1985).

Many of the earlier researchers measured statistics anxiety as a multidimensional construct using all six subscales of the STARS (Chew & Dillion, 2014). Since the development of the STARS, several researchers have examined the factor structure of the instrument and recommended that the six-factor model be revised (Hsiao, 2010; Papousek et al., 2012; Teman, 2013). In particular, Baloglu (2002) examined the construct-related validity of the six-factor model via a confirmatory factor analysis. The results of Baloglu's study revealed that the fit indices (Goodness-of-Fit Index [GFI] = .85; Comparative Fit Index [CFI] =.83; and Root Mean Square Error of Approximation [RMSEA] = .23) of the six-factor model did not fit the data well. Hsiao (2010) later compared a two-factor STARS model consisting of three subscales on each factor to the traditional one-factor STARS model with six subscales using confirmatory factor analysis. The CFI, non-normed fit index (NNFI), and RMSEA all showed acceptable values for the two-factor model (CFI = .991, NNFI = .983, and RMSEA = .060), but not for the one-factor model, thereby suggesting that the bi-dimensional model is a better and more appropriate representation of the data compared to the one-factor model. Papousek et al. (2012) further tested the factor structure of STARS using a hierarchical model with two secondary factors representing subscales more closely related to anxiety (i.e., test and class, interpretation, asking for help) and negative attitudes (computational self-concept, worth of statistics, fear of statistics teachers). Based on the Akaike's Information Criteria (AIC) and the Bayesian Information Criteria (BIC), they concluded that the hierarchical model with two correlated secondary factors performed as well as did the modified six-factor model. Although both models can be regarded as being acceptable, Papousek et al. (2012) recommended the use of the more parsimonious hierarchical model in which each student would receive two composite scores: a STARS-Anxiety score (average score from the test and class anxiety, interpretation anxiety, and fear of asking for help anxiety subscales) and a STARS-Attitudes score (average score from the computation self-concept, worth of statistics, and fear of statistics teacher subscales).

Teman (2013) further examined the structure of the STARS using several measures of fit and concluded that the six-factor model fit the data well. Additionally, Rasch modeling was used to examine whether the 51 items effectively measured various levels of the construct on each of the six dimensions of statistics anxiety (Teman, 2013). The results indicated that 20 items did not fit the Rasch model, which suggest that the 51-item STARS should be reduced to a 31-item instrument. Furthermore, a differential item functioning analysis indicated that items on each of the six subscales functioned differently for undergraduate and graduate students. This limits the ability to compare the results of the STARS across these two groups of students (Teman, 2013).

Succeeding the development of the STARS, other instruments have been developed to measure statistics anxiety. In 1991, Zeidner developed the Statistics Anxiety Inventory (SAI) to measure two dimensions of statistics anxiety: test anxiety and content anxiety. Although this instrument did produce acceptable psychometric measures, it was constructed under the assumption that mathematics and statistics anxiety are similar (Zeidner, 1991) and was modeled after the Mathematics Anxiety Scale (MARS; Richardson & Woolfolk, 1980). Similarly, Pretorius and Norman (1992) developed a unidimensional instrument to measure statistics anxiety, namely, the Statistics Anxiety Scale, which was adapted from the Mathematics Anxiety Scale (Betz, 1978). Since the development of the Statistics Anxiety Scale and the Statistics Anxiety Inventory (Zeidner, 1991), researchers have found that mathematics anxiety and statistics anxiety are related, but distinct constructs (Baloglu, 2002; Chew & Dillon, 2014). Earp (2007) later developed the Statistics Anxiety Measure, but like the STARS, the items in this instrument represented six domains, designed under the theories of statistics anxiety and attitudes towards statistics.

In 2008, Vigil-Colet, Lorenzo-Seva, and Condon developed a three-dimensional instrument in Spanish exclusively focused on statistics anxiety, the Statistical Anxiety Scale (SAS). This alternative to the STARS contains three subscales of statistics anxiety (i.e., examination anxiety, asking for help anxiety, and interpretation anxiety), with each subscale containing eight items, specifically designed for students in the social sciences. Items on this instrument were developed by a sample of university faculty members who provided situations in which students encountered during a statistics course (Vigil-Colet et al., 2008). According to the developers of the SAS, a high score on examination anxiety would imply that the student has high anxiety when taking statistics examinations. On the other hand, a high score on asking for help anxiety suggests that the student experiences high levels of anxiety when asking a teacher, peer, or tutor questions pertaining to statistics. A high score in interpretation anxiety suggests that the student encounters statistics anxiety when they have to interpret data and to understand formulas used in statistics (Vigil-Colet et al., 2008).

Evidence of construct-related validity for the Spanish version of the SAS was reported by Vigil-Colet et al. (2008) via a factor analysis using 159 participants. The three factors were reported to be correlated, which implies that the three factors are related subscales of statistics anxiety. The reported factor structure, as measured by coefficient alpha, had high score reliability values for the total scale ($\alpha = .91$), as well as for the examination anxiety subscale ($\alpha = .87$), the asking for help anxiety subscale ($\alpha = .92$), and the interpretation anxiety subscale ($\alpha = .82$), which suggest a high internal consistency of the scores. Additionally, the three factors explained 70.9% of the total common variance (Vigil-Colet et al., 2008).

Chiesi, Primi, and Carmona (2011) further confirmed the factor structure of the Italian version of the SAS. Like past results, the coefficient alpha pertaining to scores yielded by the SAS using an Italian sample was .90, suggesting high internal consistency (Chiesi, Primi, & Carmona, 2011). Additionally, Chiesi et al. (2011) conducted a cross-country validation using multi-group confirmatory factor analysis. The results confirmed that there were no statistically significant differences in the factor structure across the versions. Although the factor structure of the Italian and Spanish versions of the SAS have been validated, the English version of the SAS has yet to be validated (Chiesi et al., 2011).

Purpose of the Study

The purpose of this study was to determine the psychometric properties of the English version of the SAS scores and their relationship with scores from the Attitudes Toward Statistics Scale (ATS) using a confirmatory factor analysis. The following questions were answered:

(1) What is the factor structure of the English version of the SAS?

(2) What is the relationship between scores on items designed to measure statistics anxiety and those designed to measure attitudes towards statistics among undergraduate social science majors?

(3) To what extent does attitudes towards statistics predict statistics anxiety among undergraduate social science majors?

Despite a large body of research on the impact of statistics anxiety on student performance and interventions aimed at reducing statistics anxiety, there is no valid assessment tool in English that exclusively assesses statistics anxiety of students. It is imperative to have a valid assessment tool to assess statistics anxiety and the impact that it has on student performance, as well as to evaluate potential interventions to reduce levels of statistics anxiety (Hanna et al., 2008). The goal of this study is to score-validate an instrument that can be used by instructors who teach statistics courses for undergraduate social science majors enrolled in social science programs. The scores of this instrument will help teachers better understand their students' anxiety so they may implement instructional strategies to improve student performance in statistics.

Method

Participants

The sample consisted of 323 students majoring in a social science field (e.g., psychology, sociology), enrolled in an undergraduate, statistics for the social sciences course in the fall semester of 2016 at a college or university in the United States. The majority of the participants were female (74.4%), and the participants' ages ranged from 18 to 63 years (M = 20.5, SD = 5.24, median = 19). The majority of the participants attended face-to-face, on-campus sessions (76.9%), whereas the remaining participants were enrolled in online courses (16.5%) or a hybrid of face-to-face and online sessions (6.6%).

Sampling Procedure

Based on the type of analysis conducted for this study, it was imperative to have a sample size of at least 300 students. A two-stage sampling procedure was conducted in this study. First, simple random sampling without replacement was used to randomly select institutions with social science programs to participate in the study. Second, network sampling was used to ask instructors of statistics for social science courses to pass along the research opportunity to their students. The goal was to recruit participants similar to those used in previous validation studies (Chiesi et al., 2011; Vigil-Colet et al., 2008) for comparison purposes.

To identify instructors of statistics for social science courses, a list of colleges and universities in the United States that offer social science programs (e.g., sociology, psychology) as a field of study was obtained from the U.S. College and University Directory (2016, August 11). As of August 11, 2016, there were 1,545 institutions in the United States that offered a social science program. Of the 1,545 institutions on the list, 300 institutions were randomly selected to compensate for institutions that may not offer a statistics course for social science at the time of the study and for instructors who may not pass along the research opportunity to their students. The randbetween function in Microsoft Excel was used to select 300 numbers from a list ranging from 1 to 1,545. The 300 numbers selected represented the institutions' positions on the list of institutions with social science programs in the United States (i.e., index number). If a number was selected more than once, the institution was listed only one time. A list of 284 colleges and universities was obtained from this process.

An online search of the 2016 fall class schedule for each of the 284 schools was conducted to identify the colleges offering a statistics course for the social sciences. When an instructor's name was listed for a course, the institution's employee directory was searched to locate the instructor's e-mail address. From this process, a list of 378 instructors from 189 randomly selected colleges and universities that offered a statistics course for the social sciences in the fall were contacted via e-mail regarding the study, along with details regarding voluntary participation, the IRB approved informed consent notice, and the link to the survey. The 378 instructors were asked to pass along the research opportunity to their students.

Students were given an 8-week window to participate in the survey. The 24-item SAS, 29-item ATS, and seven demographic questions were uploaded into Qualtrics for students to answer. Students who completed the survey were given the opportunity to enter to win a $50 Visa e-gift card for their participation. For a student to be entered in the raffle for the e-gift card, they were asked to provide their contact information (i.e., name and email) at the end of the survey, which was used only to contact the winner. Participant contact information was not used in the study and was destroyed after the winner was contacted.

Sample Size

Several sample size recommendations have been made for studies involving confirmatory factor analysis (CFA). Based on the statistical theory of CFA, Kahn (2006) suggested the use of a large sample size, well over 100 participants. Likewise, Russell (2002) found that CFA conducted on sample sizes less than 100 cases tend to have inflated chi-Square statistics. Hence, minimum sample size values should be estimated by the number of parameters in the model, similar to the N:k ratio in regression analysis (Russell, 2002). Another common suggestion is that the sample size be determined as a function of the number of variables being analyzed (Stevens, 2009). For example, Bryant and Yarnold (1995) recommended a sample size of at least 10 participants for each item and a participant-to-variable ratio of at least five. Conversely, Wolf, Harrington, Clark, and Miller (2013) Monte Carlo simulation study showed that the minimum sample size-to-item and the participant-to-variable ratio are inaccurate methods for determining sample size of confirmatory factor analysis When determining minimum sample size values for confirmatory factor analysis, Wolf et al. suggested that one take into consideration potential measurement error, effects of parameter bias, weak effects, and missing data. MacCallum, Browne, and Sugawara (1996) developed a framework for determining the minimum sample size needed to achieve adequate statistical power of a CFA hypothesis test. The minimum sample size is a function of the degrees of freedom in the model. Based on the MacCallum et al. (1996) framework, to achieve power of .80, this study would need a minimum sample size of 178 participants. The actual sample size used in this study consisted of 323 participants, which was larger than the recommended sample size (MacCallum et al., 1996).

Instruments

Participants were asked to complete background information questions, the SAS, and the ATS. The background information questions consisted of demographic items pertaining to the students' gender, race, age, college major, classification in college (e.g., freshman, sophomore), and the method of course delivery (i.e., online or on-campus).

Statistical Anxiety Scale

The English version of the SAS was translated from Spanish by Vigil-Colet et al. (2008) using the back-translation procedure. This instrument consists of 24 items on a 5-point rating scale (1 = *no anxiety* and 5 = *significant anxiety*). All items on the SAS are worded positively and designed to describe common situations experienced by students enrolled in a statistics course. Additionally, the SAS was intended to measure three distinct dimensions of statistics anxiety (examination anxiety, asking for help anxiety, and interpretation anxiety) using eight items per dimension.

Attitudes Towards Statistics Scale

The ATS is a 29-item instrument designed to measure the attitudes held by college students toward statistics (Wise, 1985). Student attitudes toward statistics are evaluated on two scales: (a) attitudes towards field, which measures their attitudes toward the field of statistics via 20 items; and (b) attitudes towards course, which measures their attitudes toward the course via nine items. Each item represents a 5-point Likert type scale. The attitudes towards field subscale contains 14 positively worded items and six reverse-coded items. All nine items on the attitudes towards course subscale are reverse-coded. To score the ATS, the item scores from each of the subscales are added to obtain a subscale score and all the items are added to attain a total score. Wise (1985) reported internal consistencies of the subscales of .92 and .90 for attitudes towards field and attitudes towards course, respectively. Further, Wise (1985) documented 2-week test-retest reliabilities for attitudes towards field and attitudes towards course subscales of .82 and.91, respectively. Additionally, a factor analysis revealed that the subscales accounted for 49% of the total variance (Wise, 1985).

Pilot-Test

A pilot test was conducted using a convenience sample of 12 graduate students. This pilot test was used to determine the clarity of the items in the English version of the SAS translated by Vigil-Colet et al. (2008). Based on the feedback from the students, 10 items were revised for clarification. Details of these revisions are presented in Table 1. Many of the revisions were as small as changing one word; for instance, in Item 9, the word *doing* was changed to *completing*. Some of the revisions included adding words to the original item for detail. For example, in Item

5, the word *tutor* was added because a private teacher is referred to as a tutor in the United States. Likewise, the words *statistical software (i.e., SAS, SPSS, STATA, R)* were added to Item 17 to explain the type of printout. Due to the small sample size, statistical analysis was not conducted on these data.

Table 1

Revision of Items of the SAS

Item Number	Original Item	Revised Item
5	Asking a private teacher to explain a topic that I have not understood at all.	Asking a private teacher (tutor) to explain a topic that I do not understand at all.
9	Doing the final examination in a statistics course.	Completing the final examination in a statistics course.
12	Asking the teacher how to do an exercise.	Asking the teacher how to do a statistics problem.
13	Getting to the day before an exam without having had time to revise the syllabus.	Getting to the day before an exam and realizing that I have not prepared for a particular statistics problem.
15	Realizing, just before you go into the exam, that I have not prepared a particular exercise.	Realizing, just before you go into the exam that you have not prepared for a particular statistics problem.
16	Copying a mathematical demonstration from the blackboard while the teacher is explaining it.	Copying a mathematical demonstration from the whiteboard while the teacher is explaining it.
17	Asking one of your teachers for help in understanding a printout.	Asking one of your teachers for help in understanding a statistical software (i.e. SAS, SPSS, STATA, R) printout.
19	Seeing a classmate carefully studying the results table of a problem he has solved.	Observing a classmate carefully studying the results of a problem that they have solved.
20	Going to a statistics exam without having had enough time to revise.	Going to a statistics exam without having enough time to study.
24	Asking a private teacher to tell me how to do an exercise.	Asking a private tutor to tell me how to do an exercise.

Internal Consistency

The internal consistency score reliability for the SAS and ATS was calculated using SPSS via coefficient alpha. When a coefficient alpha is used as a measure of internal consistency, the ratio of explained variance to total variance takes into account the intercorrelation of items with the assumption that as the correlation of items increases, the value of alpha increases (Henson, 2001). Henson (2001) suggested when examining scores from instruments that contain scales for different constructs, the internal consistency should be measured for each dimension. Hence, the internal consistency of each dimension of the SAS and the ATS was measured.

Structural Validity

LISREL 9.1 was used to conduct a confirmatory factor analysis (CFA) to examine the structural validity of the SAS instrument. CFA is a statistical modeling technique of the structural equation model (SEM) family that is used to test a theory-based model of measured variables that indicate unobserved latent variables (Flora & Curran, 2004; Lei & Wu, 2007; Stevens, 2009). In an effort to replicate the methodology used in previous validation studies of the SAS (Chiesi et al., 2011), CFA was the preferred method of analysis. CFA allows the researcher to evaluate how well the observed scores of the items are indicators for the latent (unobserved) constructs (Kieffer, 1999). Diagonally weighted least squares (DWLS) method was used to estimate the model parameters. DWLS is specifically designed for ordinal data (Li, 2016). When conducting a DWLS estimation using LISREL, data can be inputted for estimation by way of the asymptotic variance or the full weight matrix (DiStefano & Morgan, 2014). To allow for the most unrestricted information (Jöreskog & Sörbom, 1996), the full weight matrix was used to estimate the parameters in this study. The goodness of overall model fit was determined by the Satorra-Bentler chi-square (χ^2_{SB}) which is a chi-square test that is corrected for non-normality. Additionally, RMSEA, CFI, and standardized root mean square residual (SRMR) were used to evaluate the goodness of overall model fit. Non-statistically significant χ^2_{SB} results indicate that there are no statistically significant differences between the two covariance matrices and that the theoretical model statistically significantly reproduces the sample covariance matrix (Schumacker & Lomax, 2010). The RMSEA fit statistic informs how well the theorized model would fit the population's covariance matrix. Browne and Cudeck (1992) suggested that RMSEA values less than .08 indicate a reasonable model fit and values below .05 indicate a good fit. The CFI fit statistic is an incremental fit index that compares the sample covariance matrix with an independence model, a model with uncorrelated latent variables (Hooper, Coughlan, & Mullen, 2008; Kline, 2011). Although CFI values close to 1 indicate good model fit, Hu and Bentler (1999) suggested a minimum CFI value of .90 to ensure misspecified theorized models are not accepted. The SRMR fit index is the difference between the sample covariance matrix and the model-implied covariance matrix; these values should be small (Kline, 2011). The suggested threshold for acceptable fit is a SRMR value less than .08 (Hu & Bentler, 1999).

Convergent and Discriminant Validity

In addition to examining the structural validity, the convergent validity and discriminant validity of the SAS was assessed. This was undertaken to ensure that scores on items from the same factor were highly correlated with each other and that scores on items from different factors were not highly correlated with each other, respectively (Onwuegbuzie, Daniel, &

Collins, 2009). This type of validity was examined to ensure that the items on the SAS measure subscales exclusively related to statistics anxiety as designed (Vigil-Colet et al., 2008).

Results

Missing Data

In this study, there were 60 cases with incomplete data. This represents 18.6% of the sample, which would be a large proportion of data to dismiss. Thus, multiple imputation (MI) estimation was used to adjust for the missing data. MI uses an iterative process of creating several imputed data sets, each containing different estimates of the missing data that are analyzed, and averaging the parameter estimates to produce one result (Peugh & Enders, 2004). This estimation method was selected because it allows the use of all the data in a dataset, and it considers the variability of the dataset even if some of the cases are incomplete (Peugh & Enders, 2004).

Reliability

Cronbach's coefficient alpha was used as the estimate of reliability of the SAS and ATS scores. Following Henson's (2001) recommendation for examining scores from instruments that contain scales for different constructs, the subscales of the SAS and the ATS were measured for internal consistency. Table 2 displays the internal consistencies of the ATS and SAS by way of Cronbach's coefficient alpha.

Table 2

Internal Consistency [and 95% Confidence Intervals] of the ATS and SAS

	SAS (24 items)				ATS (29 items)		
	Examination Anxiety	Help Anxiety	Interpretation Anxiety	Total	Attitudes Towards Field	Attitudes Towards Course	Total
Reliability	.904	.924	.823	.929	.913	.853	.920
	[.887, .920]	[.910, .937]	[.791, .852]	[.916, .941]	[.898, .927]	[.827, .876]	[.906, .933]

Note: SAS = Statistical Anxiety Scale; ATS = Attitudes Toward Statistics Scale.

The recommended benchmark value for coefficient alpha is .70 or higher for research in its early stages of development, .80 or higher for experimental treatments, and .90 as a minimum for clinical research (Nunnally, 1978; Nunnally & Bernstein, 1994). Based on the results shown in Table 2, scores from all the scales and subscales yielded high internal consistency.

Structural Validity

The goodness of fit indices for the 24-item SAS and the null model, a single-group CFA, are summarized in Table 3. The one-factor model served as a baseline for model-fit comparison with the three-factor model proposed by Vigil-Colet et al. (2008). As anticipated, the one-factor model had a poor fit to the data; this was indicated by a statistically significant χ^2_{SB} probability value, a high RMSEA and SRMR value, and a low CFI.

Table 3

Fit Statistics of the Statistical Anxiety Scale (SAS)

Model	χ^2_{SB} (*df*)	*p*	**RMSEA**	**CFI**	SRMR
Model 1: One-factor model	254.78(48.76)	<.001	.171	.577	.138
Model 2: Three-factor model	153.46(71.12)	<.001	.106	.838	.073
Model 3: Modified Two-factor model	49.37(38.13)	.105	.076	.959	.035

Note: χ^2_{SB} = Satorra-Bentler Chi-Square corrected for non-normality; *df* = degrees of freedom; *p* = probability value of χ^2_{SB}; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; SRMR = Standardized Root Mean Square Residual.

The three-factor model proposed by Vigil-Colet et al. (2008) had substantially better fit than did the one-factor model, but had a poor fit to the data. The statistically significant χ^2_{SB} probability value, a high RMSEA and SRMR value, and a low CFI indicated model misspecification for the data. Kline (2011) suggested inspecting the pattern of correlation residuals when model misspecification is present and paying close attention to absolute values greater than .10. Based on the correlation residuals and content of the instrument, the two-factor model was created. This model was created by removing Item 15 ("Realizing, just before you go into the exam that you have not prepared for a particular statistics problem") from the examination anxiety subscale, removing Item 24 ("Asking a private tutor to tell me how to do an exercise") from the fear of asking for help anxiety subscale, and removing the interpretation anxiety subscale. Items 15 and 24 were removed because their residuals were correlated with the residuals of one or more items on the same factor. The residual of Item 15 was correlated with the residuals of Item 13 ($r = .162$) and Item 20 ($r = .142$), respectively. The wording of these three items was similar; they each discussed not being prepared for a statistics examination, which indicated possible content redundancy. Item 15 had the lowest factor pattern coefficient of the three items, which suggests the removal of this item would results in loss of the least amount of information. Likewise, the correlation between the residuals of Item 24 and Item 5 was .150. The two items had similar phrasing (i.e., asking a private tutor for help with a statistics topic/problem), and Item 24 had the lowest factor pattern coefficient. The residual correlations for seven of the eight items in the interpretation anxiety subscale were larger than the suggested threshold of .10 in absolute value. Additionally, these seven items had high residual correlations with items from the other two factors. After careful examination of the wording of the questions, I decided that these seven items were not related to the other two constructs being measured. Thus, this factor was not retained.

In addition to the correlation residuals, the modification indexes were examined to identify potential adjustments to improve the fit of the two-factor model. The largest modification indexes for the two-factor model were among Items 1, 4, 13, and 20. Based on the similarity in the concept of the four items, the errors among these items were allowed to correlate. The two-factor model had an acceptable fit (χ^2_{SB}= 49.37, *df* = 38.13, *p* = .105, CFI =

.959, SRMR = .035, RMSEA = .076). This model fit the data substantially better than did the proposed three-factor model. Additionally, the standardized residuals of the two-factor model were somewhat symmetric around zero with values ranging from -1.677 to 1.945. When sample sizes are large, the standardized residuals can be used as a z test to determine whether the population covariance matrix residual is zero (Kline, 2011). Using critical values of -1.96 and 1.96, the standardized residuals confirmed that the theoretical model significantly reproduced the sample covariance matrix, indicating a good model fit.

The standardized estimates of the two-factor model are presented in Table 4 and Figure 1. The standardized factor pattern coefficients for all the items ranged from .665 to .874. The correlation between the two factors (examination anxiety and asking for help anxiety) was .588.

Table 4

Factor Pattern and Structure Coefficients (r_s) for the Two-Factor Model

	Examination Anxiety		**Asking for Help Anxiety**		
	Factor Pattern	r_s	Factor Pattern	r_s	R^2
Item **1**	.680	.680	.000	.400	.462
Item **4**	.685	.685	.000	.403	.469
Item **9**	.843	.843	.000	.496	.711
Item **11**	.833	.833	.000	.490	.694
Item **13**	.689	.689	.000	.405	.475
Item **14**	.824	.824	.000	.485	.679
Item **20**	.665	.665	.000	.391	.442
Item **3**	.000	.482	.820	.820	.672
Item **5**	.000	.439	.746	.746	.557
Item **7**	.000	.496	.844	.844	.712
Item **12**	.000	.514	.874	.874	.764
Item **17**	.000	.466	.793	.793	.629
Item **21**	.000	.505	.860	.860	.740
Item **23**	.000	.493	.839	.839	.704

Note: All standardized factor pattern are statistically significant at $p < .05$.

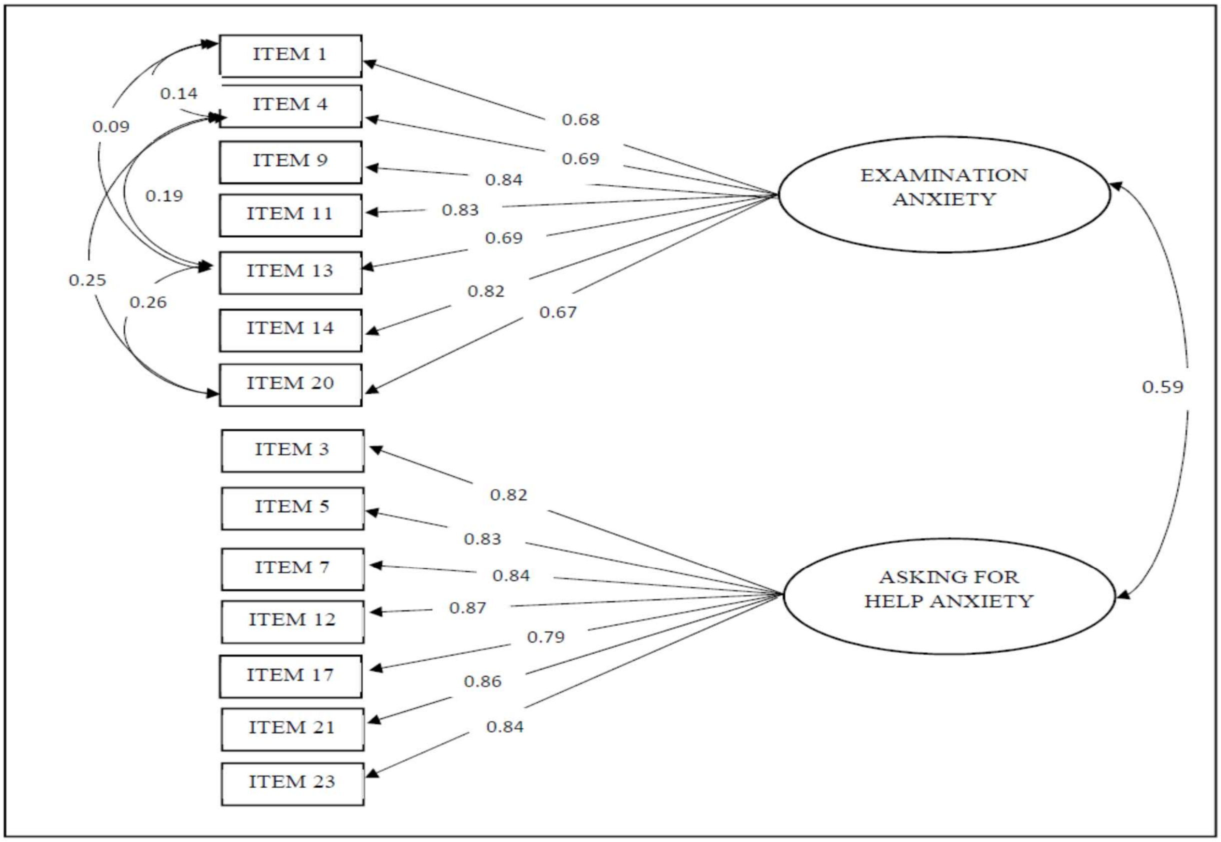

Figure 1. Standardized estimates of the two-factor model of the Statistical Anxiety Scale.

Relationship to Other Variables

Convergent Validity

To examine the convergent validity of the SAS, Pearson's product-moment correlation of the scores of the SAS subscales and the scores of the ATS subscales were evaluated. The correlations are shown in Table 5. The two subscales of the SAS and the two subscales of the ATS had moderate positive correlations ($r = .586$ *and* $r = .590$, respectively). As expected, examination anxiety had a moderately negative correlation with attitudes towards course ($r = -.575$). Hence, evidence of convergent validity is supported by the moderate correlations between the variables that were expected to have a relationship (i.e., examination anxiety with asking for help anxiety, examination anxiety with attitudes towards course, attitudes towards course with attitudes towards field).

Table 5

Correlations between the SAS and ATS Subscales

	SAS		ATS	
	Examination Anxiety	Asking for Help Anxiety	Attitudes towards Field	Attitudes towards Course
Asking for Help Anxiety	.586*			
Attitudes towards Field	-.028	-.112		
Attitudes towards Course	-.575*	-.387	.590*	

Note: SAS = Statistical Anxiety Scale; ATS = Attitudes Toward Statistics Scale. *$p < .05$.

Discriminant validity

To establish discriminant validity of the SAS, the average variance extracted (AVE) from each factor was compared to the shared variance of the factors. Farrell (2010) describes discriminant validity as the extent to which factors measuring the same construct account for more shared variance than do factors not theoretically associated with the construct. Thus, if the AVE of two factors is greater than the shared variance between the same two factors, discriminant validity is supported.

Table 6

Shared Variance and Average Variance Extracted from the factors

	SAS		ATS	
	Examination Anxiety	Asking for Help Anxiety	Attitudes towards Field	Attitudes towards Course
Asking for Help Anxiety	.343			
Attitudes towards Field	.001	.013		
Attitudes towards Course	.331	.150	.348	
AVE	.092	.110	.203	.096

Note. AVE = average variance extracted per factor.

The shared variance and AVE of the factors shown in Table 6 indicate some evidence of discriminant validity between examination anxiety and attitudes towards field as well as between asking for help anxiety and attitudes towards field. Likewise, the correlations between the SAS factors and the ATS factors support discriminant validity (see Table 5). Shaffer, DeGeest, and Li (2016) describe discriminant validity as evidence that theoretically distinct factors are not highly correlated.

Attitudes as a Predictor of Anxiety

SEM was used to analyze the structural regression model shown in Figure 2 to determine whether attitudes towards course and attitudes towards the field were predictors of examination anxiety and asking for help anxiety. The SEM estimates are shown in Table 7. Attitudes towards field and attitudes towards course together explained 28.5% of the variation in students' examination anxiety scores. Attitudes towards field and attitudes towards course together explained 11.6% of the variation in students' asking for help anxiety scores. An evaluation of the structure coefficients and regression beta weights showed that attitudes towards course attributed to the largest amount of variation in examination anxiety. Attitudes towards course also had the largest contribution to predicting the variation in asking for help anxiety.

Table 7

Regression Parameter Estimates

Paths	Coefficient	β	Std. Error	*t*	*p*	r_s
Attitudes towards course -> Examination Anxiety	-.606	-.605	.054	-11.318	<.001	-.873
Attitudes towards field ->Examination Anxiety	.158	.296	.029	5.536	<.001	.021
Attitudes towards course -> Asking for Help Anxiety	-.390	-.375	.062	-6.305	<.001	-.966
Attitudes towards field ->Asking for Help Anxiety	.054	.098	.033	1.642	.101	-.232

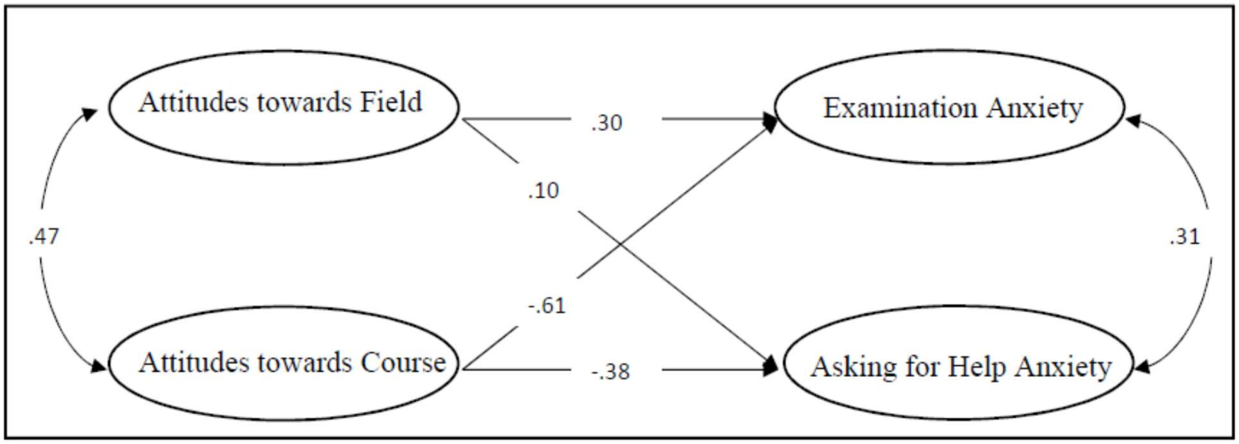

Figure 2. Structural regression model for predicting statistics anxiety.

Discussion

The purpose of this study was to test the psychometric properties of the English version of the Statistical Anxiety Scale (Vigil-Colet et al., 2008), a 24-item instrument designed exclusively to measure statistics anxiety. The overall instrument and subscale levels had high internal consistencies (α = .823 to α = .929), consistent with those reported in previous validation studies (Chiesi et al., 2011; Vigil-Colet et al., 2008). Although the internal consistency of the three-factor model was high, the structure of this model lacked an acceptable model fit for the current sample. After an examination of the correlations between the residuals and the content of the items, a revised two-factor model was proposed and tested. The structure of the two-factor model had an acceptable fit for the current sample (χ^2_{SB}= 49.37, *df* = 38.13, *p* = .105, CFI = .959, SRMR = .035, RMSEA = .076). Additionally, all items loaded statistically significantly onto their theorized factor and there was a moderate correlation between the two factors (*r* = .588).

The two-factor model was developed by removing one item from the asking for help anxiety subscale and one item from the examination anxiety subscale. The interpretation anxiety subscale was removed from the model due to seven of the eight items having high residual correlations with items from the other two subscales. Likewise, the content of the seven items on the interpretation scale were not related to the subject matter of the other two subscales. One potential reason the items on the interpretation scale contributed to the unacceptable model fit might be the methods used in the development of the SAS. During the development of instrument, Vigil-Colet et al. (2008) adapted five of the eight items on the interpretation scale from the STARS. During a Rasch item analysis of the STARS, Teman (2013) suggested the removal of two of the items that Vigil-Colet et al. (2008) used on the interpretation subscale due to item misfit. While all the items on the interpretation scale examined by Teman (2013) were not identical to all the items on the interpretation scale of the SAS, the results of this study still indicate potential validity issues with two of the items selected by Vigil-Colet et al. (2008) to measure the construct.

Convergent and discriminant validity of the two-factor SAS was examined to ensure the factors of the SAS measured the intended construct—statistics anxiety. This was undertaken to ensure that scores on items from theoretically similar constructs are highly correlated with each other and that scores on items from theoretically distinct constructs are not highly correlated (Onwuegbuzie, Daniel, & Collins, 2009; Shaffer, DeGeest, & Li, 2016). As expected, the correlations between the two factors of the SAS and the two factors of the ATS provided evidence of convergent and discriminant validity. The moderately negative correlation between examination anxiety and attitudes towards course ($r = -.575$) was consistent with previous studies that considered the relationship between statistics anxiety and attitudes towards statistics (Chiesi & Primi, 2010; Mji & Onwuegbuzie, 2004). Furthermore, this negative correlation supports Chew and Dillon's (2014) definition of statistics anxiety; a negative emotional state stimulated from any form of interaction with statistics and exacerbated by negative attitudes towards it.

In addition to examining the structural validity of the SAS and its relationship to attitudes towards statistics, a structural regression model was used to examine the extent to which attitudes towards the statistics course and attitudes towards the field of statistics predict examination anxiety and asking for help anxiety. Both subscales of attitudes explained 28.5% of the variation in examination anxiety with attitudes towards the course having the largest contribution towards examination anxiety (r_s = -.87). Additionally, the two subscales of attitudes towards statistics together explained 11.6% of the variation in asking for help anxiety scores, with attitudes towards course being the only statistically significant predictor of asking for help anxiety. This suggests that attitudes towards course and attitudes towards the field moderately predicted examination anxiety, with attitudes towards the course having a stronger influence. Likewise, attitudes towards course moderately predicted asking for help anxiety.

The results of this study contribute to the field of statistics anxiety by providing an instrument exclusively designed to measure statistics anxiety. Currently, there is limited research that distinguishes statistics anxiety from related variables such as attitudes towards statistics (Chew & Dillon, 2014). Moreover, the two terms are often used interchangeably in research (Nasser, 2004). This might be due to the use of the STARS as a multidimensional measure of statistics anxiety in many of the previous studies on statistics anxiety. Since the development of the STARS, several researchers have examined the factor structure of the instrument and recommended that the six-factor model be revised (Hsiao, 2010; Papousek et al., 2012; Teman, 2013). Furthermore, the STARS were found to measure two constructs, statistics anxiety and attitudes towards statistics (Hsiao, 2010; Papousek et al., 2012). Furthering the need to analyze the results of STARS as two composite scores: a STARS-Anxiety score and a STARS-Attitudes score (Papousek et al., 2012). Earlier research used all six subscales of the STARS as a measure of statistics anxiety; hence examining both statistics anxiety and attitudes towards statistics as if they were one construct (Chew & Dillon, 2014). As researchers continue to evaluate interventions aimed at reducing statistics anxiety and the impact statistics anxiety has on students learning, it is important to have a valid tool to assess the extent to which students are experiencing statistics anxiety (Hanna et al., 2008). The validation of the revised two-factor SAS provides researchers with an instrument designed to solely measure statistics anxiety. It is recommended that future research continues to validate the revised two-factor SAS using various samples (e.g., graduate students).

Limitations

As in any study, there were limitations that might have influenced the results of the current study. Although random sampling was used to identify participants, the sample was limited to professors who passed along the research opportunity to their students. Additionally, the background of the students (i.e., previous enrollment in statistics courses, reasons for enrolling in course) and students with majors in the pure sciences (i.e., chemistry, biology, mathematics, etc.) were not considered in the study. It is recommended that future studies examine students from other academic fields enrolled in statistics courses designed for the social sciences. Also, future studies should examine the relationship among attitudes towards statistics, statistics anxiety, and other variables such as academic achievement and instructional style.

References

Baloglu, M. (2002). Psychometric properties of the statistics anxiety rating scale. *Psychological Report*, 90, 315-325. doi:10.2466/pr0.2002.90.1.315

Baloglu, M. (2003). Individual differences in statistics anxiety among college students. *Personality and Individual Differences*, *34*, 855-865. doi:10.1016/S01918869(02)00076-4

Betz, N. E. (1978). Prevalence, distribution, and correlates of math anxiety in college students. *Journal of Counseling Psychology*, *25*, 441-448. doi:10.1037/0022-0167.25.5.441

Birenbaum, M., & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, *36*, 93-98. doi:10.1080/0013188940360110

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230-258. doi:10.1177/0049124192021002005

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington DC: American Psychological Association.

Chew, P. K. H., & Dillon, D. B. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science*, *9*, 196-208. doi:10.1177/1745691613518077

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal*, *9*, 6-26.

Chiesi, F., Primi, C., & Carmona, J. (2011). Measuring statistics anxiety: Cross-country validity of the statistical anxiety scale (SAS). *Journal of Psychoeducational Assessment*, *29*, 559-569. doi:10.1177/0734282911404985

Cruise, R. J., & Wilkins, E. M. (1980). *STARS: Statistical Anxiety Rating Scale*. Unpublished manuscript, Andrews University, Berrien Springs, MI.

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, Las Vegas, NV*.

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*, 425-438. doi:10.1080/10705511.2014.915373

Earp, M. S. (2007). *Development and validation of the statistics anxiety measure* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global. (UMI No. 3279112).

Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research*, *63*, 324-327. doi:10.1016/j.jbusres.2009.05.003

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*, 466-491. doi:10.1037/1082-989X.9.4.466

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 Individuals Differences*, *45*, 68-74. doi:10.1016/j.paid.2008.02.021

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha (Methods, plainly speaking). *Measurement and Evaluation in Counseling and Development, 34*, 177-189.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53-60.

Hsiao, T. (2010). The statistical anxiety rating scale: Further evidence for multidimensionality. *Psychological Reports*, *107*, 977-982. doi:10.2466/07.11.PR0.107.6.977-982

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equational Modeling*, *6*, 1-55. doi:10.1080/10705519909540118

Jöreskog, K.G., & Sörbom, D. (1996). *LISREL user's reference guide*. Chicago, IL: Scientific Software International.

Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist*, *34*, 684-718. doi:10.1177/0011000006286347

Kieffer, K. M. (1999). An introductory primer on the appropriate use of exploratory and confirmatory factor analysis. *Research in Schools*, *6*(2), 75-92.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.

Lei, P., & Wu, Q. (2007). An NCME instructional module on introduction to structural equation modeling: issues and practical considerations. *Educational Measurement: Issues and Practice*, *26*(3), 33-43. doi:10.1111/j/1745-3992.2007.00099.x

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonal weighted least squares. *Behavior Research Methods*, *48*, 936-949. doi:10.3758/s13428-015-0619-7

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1, 10-149. doi:10.1037/1082-989X.1.2.130

Mji, A., & Onwuegbuzie, A. J. (2004). Evidence of score reliability and validity of the statistical anxiety rating scale among technikon students in South Africa. *Measurement and Evaluation in Counseling and Development*, *36*, 238-251.

Nasser, F. M. (2004). Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking pre-service teachers in introductory statistics. *Journal of Statistics Education*, 12(*1*). Retrieved from https://ww2.amstat.org/publications/jse/v12n1/nasser.html

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Nunnally, J.C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Onwuegbuzie, A. J. (1997). Writing a research proposal: The role of library anxiety, statistics anxiety, and composition anxiety. *Library & Information Science Research*, *19*, 5-33. doi:10.1016/S0740-8188(97)90003-7

Onwuegbuzie, A. J. (1999). Statistics anxiety among African American graduate students: An affective filter? *Journal of Black Psychology*, *25*, 189-209. doi:10.1177/0095798499025002004

Onwuegbuzie, A. J. (2000). Statistics anxiety and the role of self-perceptions. *The Journal of Educational Research*, *93*, 323-330. doi:10.1080/00220670009598724

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

Onwuegbuzie, A. J., Da Ros, D., & Ryan, J. M. (1997). The components of statistics anxiety: A phenomenological study. *Focus on Learning Problems in Mathematics*, *19*(4), 11-35.

Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. T (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity*, *43*, 197-209. doi:10.1007/s11135-007-9112-4

Onwuegbuzie, A. J., Leech, N. L., Murtonen, M., & Tahtinen, 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

Papousek, I., Ruggeri, K., Macher, D., Paechter, M., Heene, M., Weiss, E. M., & Freudenthaler, H. H. (2012). Psychometric evaluation and experimental validation of the statistics anxiety rating scale. *Journal of Personality Assessment, 94*, 82-91. doi:10.1080/00223891.2011.627959

Pan, W., & Tang, M. (2004). Examining the effectiveness of innovative instructional methods on reducing statistics anxiety for graduate students in the social sciences. *Journal of Instructional Psychology*, *31*, 149-159.

Pretorius, T. B., & Norman, A. M. (1992). Psychometric data on the statistics anxiety scale for a sample of South African students. *Educational and Psychological Measurement*, *52*, 933-937. doi:10.1177/0013164492052004015

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, *74*, 525-556. doi:10.3102/00346543074004525

Richardson, F. C., & Woolfork, R. (1980). Mathematics anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 271-288). Hillsdale, NJ: Erlbaum.

Russell, D. W. (2002). In search of underlying dimensions: the use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, 28, 1692-1646. doi:10.1177/014616702237645

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Taylor & Francis.

Shaffer, J.A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, *19*, 80-110. doi:10.1177/1094428115598239

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. New York, NY: Taylor & Francis.

Teman, E. D. (2013). A Rasch analysis of the statistical anxiety rating scale. *Journal of Applied Measurement*, *14*, 414-434.

US College and University Directory (2016, August 11). *UnivSearch.com*. Retrieved from http://www.univsearch.com/family-search.php?family=45

Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L. (2008). Development and validation of the Statistical Anxiety Scale. *Psicothema*, *20*(1), 174-180.

Williams, A. S. (2013). Worry, intolerance of uncertainty, and statistics anxiety. *Statistics Education Research Journal*, *12*, 48-59.

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, *45*, 401-405.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*, 913-934. doi:10.1177/0013164413495237

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, *61*, 319-328. doi:10.1111/j.2044-8279.1991.tb00989.x

APPENDIX A

EXTENDED LITERATURE REVIEW

Statistics Anxiety and Individual Differences

Students with statistics anxiety have been found to show physiological signs such as increased heart rate, perspiration, dry mouth, headaches, muscle pain, sick feelings, and reduced sex drive (Onwuegbuzie et al., 1997). To understand better how statistics anxiety affects individuals, many researchers have explored the relationship between individual differences and statistics anxiety. For example, using a multivariate approach, Baloglu (2003) examined the relationship between college students' gender, age, and statistics anxiety levels while accounting for differences in previous mathematics experience. The results showed that age had a statistically significant effect on statistics anxiety, but not gender. Students older than 27 years of age had higher levels of statistics-test and-class anxiety compared to students younger than 21 years of age. However, students younger than 21 years of age had the lowest score on the worth of statistics subscale, whereas students older than 27 years of age scored highest.

Pan and Tang (2004) further examined the relationship between individual differences and statistics anxiety using an analysis of covariance (ANCOVA) model. The results suggested that a student's age, the number of mathematics or statistics courses previously taken, and experience in academic research had significant effects on statistics anxiety. This implies that an older student with little research experience who has taken few mathematics or statistics courses likely will exhibit higher levels of statistics anxiety than will a younger student who has taken several mathematics or statistics courses and has research experience. Roberte-Luna and Sherry (2008) extended the research on statistics anxiety and gender by incorporating cognitive learning strategies (i.e., procrastination, time and study environment, rehearsal, elaboration, organization, critical thinking, metacognitive self-regulation, effort regulation, peer learning, and help seeking) using descriptive discriminant analysis and canonical correlation analysis. Statistically significant differences between men and women on subscales of statistics anxiety and learning strategies were identified, but these differences represented a small effect size, which indicated that the gender differences are relatively minute and weak. It should be noted that differences existed in men's and women's relationship between learning strategies and statistics anxiety. In particular, procrastination and organization were strongly related to statistics anxiety in men, whereas in women, there was a relationship among all the learning strategies, except peer learning and statistics anxiety (Roberte-Luna & Sherry, 2008).

Bell (2003) extended the research on individual differences and statistics anxiety by examining the differences in statistics anxiety between non-traditional and traditional college students. The results revealed that non-traditional students experienced statistically significant higher levels of test and class anxiety than did traditional students. Additionally, there was a statistically significant negative relationship between non-traditional students' fear of statistics teacher and final grades. This indicates that final grades decrease with increases in anxiety (Bell, 2003). Bell (2008) further examined the differences in statistics anxiety between international and domestic students enrolled in a business statistics course. The findings indicated that international students experienced higher levels of statistics anxiety, particularly interpretation anxiety, computational self-concept, and fear of asking for help. Similarly, African American students have been found to experience statistically significantly higher levels of statistics anxiety compared to White American students, specifically in the areas of worth of statistics, interpretation anxiety, and test and class anxiety (Onwuegbuzie, 1999).

In earlier studies, statistics anxiety was compared between students who had previous statistics training and those who had not. The results indicated that previous statistics training did not affect statistics anxiety, nor was statistics anxiety related to students' grades (Birenbaum &

Eylath, 1994). In addition, inductive reasoning ability was negatively related to statistics anxiety and positively related to statistics course grades for all students. Dykeman (2011) added to the research on individual differences by exploring the anxiety levels of college students enrolled in statistics courses compared to those enrolled in other education courses. Results showed that students enrolled in statistics courses experienced higher levels of anxiety compared to students enrolled in non-statistics courses. Furthermore, students without prior experience in statistics experienced higher levels of anxiety, lower self-perceived readiness, and lower self-efficacy than did students with prior statistics experience (Dykeman, 2011).

Statistics Anxiety and Other Latent Constructs

In addition to the relationship between individual differences and statistics anxiety, a large body of research has examined the association of statistics anxiety with other latent constructs (e.g., worry, self-efficacy). For example, Williams (2013) examined the relationship among intolerance of uncertainty, worry, and statistics anxiety in graduate statistics students. The results of this study revealed that intolerance of uncertainty accounted for approximately 30% of the variation in worry with respect to statistics anxiety. Additionally, there were moderate effect sizes between worry and three subscales of the STARS (i.e., interpretation anxiety, test and class anxiety, and computation self-concept). This indicates that worry is related to statistics anxiety. Bandalos, Yates, and Thorndike-Christ (1995) further revealed that the worry component of statistics test anxiety was negatively related to self-efficacy in statistics. This implies that students with low self-efficacy in statistics tend to agonize more over their performance on statistics tests. In a later study on the relationship between self-perception and statistics anxiety, Onwuegbuzie (2000) found that perceived creativity, intellectual ability, and scholastic competence were associated with all subscales of statistics anxiety, whereas perceived job competence was not related to statistics anxiety. These findings suggest that statistics anxiety is an academic-related phenomenon and that many statistics-anxious students tend to pursue careers that require minimal quantitative skills (Onwuegbuzie, 2000).

Extending on these findings, Onwuegbuzie (2003) further examined the relationship among statistics anxiety, research anxiety, study habits, and statistics achievement. The results of this study revealed that statistics anxiety was negatively associated with statistics achievement, and expectations of statistics achievement was positively related to statistics achievement. These findings indicate that students with low expectations of their statistics ability tend to exhibit behaviors that might lead to underachievement (Onwuegbuzie, 2003). Additionally, students exhibiting low levels of hope in learning statistics tend to show higher levels of anxiety pertaining to statistics (Onwuegbuzie, 1998). Other researchers have found student interest in statistics to be negatively associated with the intensity of their statistics anxiety (Macher et al., 2013) and found the need for achievement to be unrelated to test performance and statistics anxiety (Keeley, Zayac, & Correia, 2008). In addition, graduate students' preference for numerical information had a strong negative association with Worth of Statistics ($r = -.59$) and Computational Self-Concept ($r = -.49$) (Williams, 2014). This suggests that students who enjoy working with numerical information have a high perception of the significance of statistics and a high self-perception of their ability to understand and to calculate statistics problems.

Kesici, Baloglu, and Deniz (2011) further investigated the relationship between statistics anxiety and self-regulated learning strategies such as cognitive, metacognitive, organization, and resource management strategies. They found a statistically significant relationship between self-regulated learning strategies (i.e., cognitive and metacognitive) and statistics anxiety (i.e., computational self-concept and worth of statistics). This implies that students who perceive statistics useful (i.e., worth of statistics) and have low levels of anxiety when solving statistical problems (i.e., computational self-concept) tend to use learning strategies such as rehearsal, elaboration, organization, critical thinking, metacognitive self-regulation, effort regulation, help seeking, and time management (Kesici, Baloglu, & Deniz, 2011). Additionally, several dimensions of Gardner's (1983) theory of multiple intelligence have been associated with statistics anxiety. In particular, higher levels of interpretation anxiety, test and class anxiety, computational self-concept, and worth of statistics were associated with students whose abilities were in favor of spatial and interpersonal intelligence and less in favor of linguistic and logical-mathematical intelligence (Onwuegbuzie & Daley, 1997).

Statistics anxiety also has been linked to personality factors such as perfectionism and procrastination. Specifically, Onwuegbuzie and Daley (1999) investigated the relationship between three dimensions of perfectionism and the six dimensions of statistics anxiety for a sample of students enrolled in a graduate-level research methodology course. The results revealed that statistics anxiety was predicted by other-oriented perfectionism, which refers to perfectionists who hold unrealistically high expectations of others, and socially prescribed perfectionism, which refers to perfectionists who believe that others expect them to be perfect. In particular, students with high levels of other-oriented and socially prescribed perfectionism tend to have higher levels of interpretation anxiety, computations self-concept, and fear of asking for help (Onwuegbuzie & Daley, 1999). Walsh and Ugumba-Agwunobi (2002) extended this study on statistics anxiety and perfectionism by examining the addition of trait anxiety and procrastination. They found all three components of procrastination to be statistically significant predictors of various subscales of STARS. Additionally, trait anxiety statistically significantly predicted test and class anxiety, interpretation anxiety, and fear of asking for help whereas procrastination statistically significantly predicted interpretation anxiety, fear of asking for help, and fear of statistics teacher. The results suggest that self-imposed engagement in evaluation concern, fear of failure, and need for approval are related to trait anxiety, procrastination, and socially prescribed perfectionism (Walsh & Ugumba-Agwunobi, 2002). Onwuegbuzie (2004) additionally explored the relationship between statistics anxiety and procrastination using a sample of graduate students. A key finding from this study was that procrastination produced by fear of failure and task aversion was statistically significantly related to all six dimensions of statistics anxiety (i.e., computational self-concept, fear of asking for help, fear of statistics teachers, interpretation anxiety, test and class anxiety, and worth of statistics).

Statistics Anxiety and Other Variables

In addition to the relationship between statistics anxiety and personality constructs, Bell (2001) assessed the effect of the length of the term on the anxiety level in students taking a statistics course. Using the STARS, he examined the differences in anxiety levels among students taking a statistics course during the 3-week intersession term, a 5-week summer term, and a 16-week semester term. The findings suggest that the level of statistics anxiety increased as the length of the term decreased. In particular, students enrolled in the intersession or summer term experienced higher levels of statistics anxiety in worth of statistics, interpretation anxiety, test and class anxiety, and computational self-concept (Bell, 2001).

As many universities move towards offering online degree and certification programs, the number of statistics courses offered online also might increase. Many challenges are presented when teaching and learning in an online environment; for demanding classes, such as statistics, these challenges are magnified (Dunn, 2014). DeVaney (2010) compared statistics anxiety among graduate students enrolled in online and on-campus statistics courses at the beginning and end of the course. He documented that students enrolled in the online statistics course tend to have higher levels of statistics anxiety and less positive attitudes at the beginning of the course than do those enrolled in an on-campus course. Additionally, there was no statistically significant difference in statistics anxiety between online and on-campus students at the end of the course (DeVaney, 2010). Later, Dunn (2014) explored the influence of academic self-regulation, intrinsic motivation, and statistics anxiety on passive procrastination for graduate students enrolled in online statistics and research methodology courses. The results indicated that as statistics anxiety increased, passive procrastination increased. Additionally, intrinsic motivation was found to mediate this relationship.

Statistics Anxiety and Cross-Cultural Differences

Several researchers have contributed to the literature on statistics anxiety by examining cross-cultural differences. In a comparison of Turkish and U.S. college students, Baloglu, Deniz, and Kesici (2011) found that both groups of students experienced the highest level of anxiety with respect to test and class anxiety. Although there were no statistically significant differences in test and class anxiety and computational self-concept between the two groups, U.S. students tended to believe that statistics was less useful (worth of statistics) and experience higher levels of anxiety when interpreting data (interpretation anxiety), asking for assistance pertaining to statistics (fear of asking for help), and interacting with the statistics instructor (fear of statistics teacher) (Baloglu, Deniz, & Kesici, 2011).

Using a South African population, Mji and Onwuegbuzie (2004) examined the factor structure of the STARS and found consistent results in terms of internal consistency reliability as well as attitudes towards statistics was a statistically significant predictor of five of the six subscales of STARS (interpretation anxiety was not statistically significant). Hanna et al. (2008)

also attempted to test the factor structure of the STARS using a United Kingdom sample. Their findings were consistent with previous research from students sampled in the United States (e.g., Baloglu, 2002; Cruise et al., 1985; Keeley et al., 2008), which suggests that there are no differences in the structure of STARS between students in the United Kingdom and the United States. Hanna et al. (2008) noted that many of the items on the STARS measured concepts related to statistics anxiety such as the perceived worth of statistics. In addition to a cross-country validation, one-factor, four-factor, and six-factor models were tested for reasonable fit. Based on the RMSEA-, IFIF-, CFI-, and SRMR-suggested criteria for model fit, all the indices indicated that a six-factor model was the best fit for STARS.

Strategies to Reduce Statistics Anxiety

A variety of strategies have been tested in an effort to reduce statistics anxiety among students. For instance, Pan and Tang (2004) found a reduction in student levels of statistics anxiety after implementing application-oriented teaching methods and applying attentiveness to student anxiety. Application-oriented teaching methods included journal article critiquing and biweekly essay writing pertaining to what the students learned in class, whereas the instructor attentiveness to student anxiety included an orientation letter prior to class, flexible and extra office hours, a midterm class survey, a cheat sheet to use during examination, and an optional pass/fail grading system (Pan & Tang, 2004). Another strategy used in the classroom to reduce student levels of statistics anxiety was cartoon humor. Although Schacht and Stewart (1990) could not statistically link humor to reducing statistics anxiety, the sample of students believed that the incorporation of humor in the class reduced their anxiety levels. Furthermore, the students believed that the use of humor contributed to an enjoyable learning environment, although it did not increase their levels of understanding. In a similar study, some students believed that the incorporation of humor provided a mental break and broke up the content, whereas students who were highly motivated and interested in the topic found the use of humor to be distracting and irrelevant (Neumann, Hood, & Neumann, 2009).

Bell (2003) suggested the use of teaching strategies to assist students with high levels of statistics anxiety. Such strategies include distributing solved tests (i.e., old tests with step-by-step explanations of the solution for each problem) and unsolved tests (i.e., old tests without the solutions) to counteract test and class anxiety as well as encouraging the use of conversation partners to reduce the fear of asking for help. Additionally, he encouraged professors to post solved copies of the examination after returning the graded exam to lessen the fear of asking for help. The return of graded papers and examinations within one class period also has been noted as a strength when working with students experiencing statistics anxiety (Bell, 2003). Another approach to reducing statistics anxiety is the 1-minute paper strategy (Chiou, Wang, & Lee, 2014). In this strategy, students use the last 5 to 10 minutes of class to answer the following two questions: (a) What is the most important concept you learned in class today? and (b) What questions remain unanswered? Using a quasi-experimental research design, researchers found that students exposed to the 1-minute paper strategy exhibited lower levels of statistics anxiety and higher academic achievement in statistics than did the control group (Chiou et al., 2014). Statistics instructors also can decrease student levels of statistics anxiety through exhibiting immediate behavior, verbal and non-verbal behaviors that demonstrate an enjoyable and welcoming environment for students (Williams, 2010). Using a pre-test post-test designed study, Williams (2010) found a decrease in student levels of statistics anxiety when instructors demonstrated immediate behavior. In particular, instructor immediacy explained the largest amount of variance in fear of statistics teacher (20%), computational self-concept (11%), and test and class anxiety (10%). The results from this study suggest that instructors who engage in positive communicative behaviors tend to reduce student discomfort in relating to the instructor, solving statistics problems, and taking statistics tests (Williams, 2010). Macheski, Buhrmann, Lowney, and Bush (2008) further suggested that instructors of difficult subjects such as statistics transform the traditional classroom into a community of learners in which students are actively involved in the day-to-day course dynamics, course content is used to create a common language of discourse, and the classroom serves as a supportive emotional environment.

In an attempt to reduce statistics anxiety, it has also been suggested that universities eliminate traditional statistics courses from the graduate curriculum. Traditional descriptive and inferential statistics courses would be replaced with courses that emphasize the link among research questions, research methods (i.e., qualitative, quantitative, and mixed), and data analysis techniques. Onwuegbuzie et al. (2010) suggested that the de-emphasizing of statistics in courses will lead to a reduction in student levels of statistics anxiety, because statistical analysis will represent a single step in the process of research and data analysis. Furthermore, Perepiczka, Chandler, and Becerra (2011) posited that research methods and statistics should be encompassed in every course in the graduate curriculum for continual exposure to the terminology and to increase the perceived usefulness of statistics in the students' professions.

Instruments to Measure Attitudes towards Statistics

In a review of instruments designed to measure student attitudes towards statistics, Nolan, Beran, and Heckler (2012) found that only four out of 15 survey instruments had been used in more than one peer-reviewed, published article examining score validity. The four surveys were the Statistics Attitude Survey (Roberts & Bilderback, 1980), the Attitudes Toward Statistics (ATS) scale (Wise, 1985), the Survey of Attitudes Toward Statistics – 28 (SATS-28; Schau, Stevens, Dauphinee, & Del Vecchio, 1995) and the Survey of Attitudes Toward Statistics – 36 (Schau, 2003).

The Statistics Attitude Survey is a 34-item scale designed to measure attitudes towards statistics and to predict performance in statistics as well as to serve as an aid for instructors in detecting students with a fear of statistics (Roberts & Bilderback, 1980). The design of this instrument was later criticized because many of the items measured student achievement instead of attitudes, and the items were not designed for students in the beginning of a statistics course (Wise, 1985).

As a result, Wise (1985) designed a 29-item instrument designed to measure the attitudes held by college students toward statistics, namely, the ATS. Student attitudes toward statistics are evaluated on two scales, field, which measures their attitudes toward the field of statistics via 20 items, and course, which measures their attitudes toward the course via nine items, both of which are assessed on a 5-point Likert-format scale. The internal consistency pertaining to scores generated by the instrument was reported for the subscales, Field ($\alpha = .92$) and Course ($\alpha = .90$), as well as 2-week test-retest reliabilities, $\alpha = .82$ and $\alpha = .91$, respectively (Wise, 1985). Several researchers have reported comparable internal consistency scores (Cashin & Elmore, 2005; Mji, 2009; Shultz & Koshino, 1998).

In 1995, Schau, Stevens, Dauphinee, and Del Vecchio developed a 28-item instrument, SATS-28, designed to measure four subscales of attitudes towards statistics: affect, cognitive competence, value, and difficulty. In an investigation of the construct-related validity of the SATS-28, the results from a confirmatory factor analysis indicated that the instrument was a two-factor model with two factors like those in the ATS (Cashin & Elmore, 2005). The field items from the ATS measured the same construct as did the value items from the SATS-28, whereas the course items from the ATS corresponded with the affect, cognitive competence, and difficulty items from the SATS-28.

In 2003, Schau added two dimensions, interest and effort, and eight items to the SATS-28, creating a new version of the instrument, the SATS – 36. Vanhoof, Kuppens, Sotos, Verschaffel, and Onghena (2011) investigated the structure of the SATS-36 through a confirmatory factor analysis and found that the six-factor model had acceptable model fit, but several items needed modifications. It was suggested that three items be deleted from the Difficulty subscale due to low factor loadings and that one be removed from the Affect subscale, because two items were found to be closely related. Due to the structural validity issues in the SATS-28 and SATS-36, the ATS was selected to measure attitudes towards statistics in this study.

APPENDIX B

DETAILED METHODOLOGY

Research Design

Data Cleaning

The 53 Likert-type items and seven demographic items were extracted from Qualtrics and moved into the Statistical Package for the Social Sciences (SPSS) for initial data cleaning. The data-cleaning process included removing any recorded cases that were missing all the responses from the dataset, ensuring that reverse-scored items were truly reversed in the dataset, confirming that missing responses were coded as missing in the dataset, and removing any variables created by Qualtrics that were not needed in the dataset (e.g., start and end time of survey, respondent ID). Additionally, SPSS was used to examine outliers.

Outliers

The Mahalanobis distance (D) was used to assess outliers in the data. D is the distance, in standard deviation units, between each participant score and the sample means of all variables (i.e., items on the survey). The squared value of D is distributed as a chi-square statistic with degrees of freedom equal to the number of variables for large samples (Kline, 2011). To reduce the likelihood of a Type I error, Kline (2011) recommended the use of a conservative level of statistical significance, such as $p < .001$. Using this recommendation, there were five cases with *p*-values below .001. These cases were considered extreme scores or outliers and were removed from the dataset.

Reliability

The internal consistency of the reliability of the scores for the SAS and ATS was calculated using SPSS by way of coefficient alpha. When a coefficient alpha is used as a measure of internal consistency, the ratio of explained variance to total variance takes into account the intercorrelation of items with the assumption that as the correlation of items increases, the value of alpha increases (Henson, 2001). Henson (2001) suggested that when examining scores from instruments that contain scales for different constructs, the internal consistency should be measured for each dimension. Hence, the internal consistency of each dimension of the SAS and the ATS was measured.

The recommended benchmark value for reliability of the scores by way of coefficient alpha is .70 or higher for research in its early stages of development, .80 or higher for experimental treatments, and .90 as a minimum for clinical research (Nunnally, 1978; Nunnally & Bernstein, 1994). Clark and Watson (1995) suggested that once the benchmark value is achieved, one should focus their attention on maximizing validity and less on increasing reliability. Hence, the focus of this study is on maximizing validity of the instrument by way of confirmatory factor analysis (CFA).

Structural Validity

LISREL 9.1 was used to conduct a CFA to examine the structural validity of the SAS instrument. CFA is a statistical modeling technique of the structural equation model (SEM) family that is used to test a theory-based model of measured variables that indicate or infer unobserved latent variables (Flora & Curran, 2004; Lei & Wu, 2007; Stevens, 2009). CFA was selected in this study because it allows one to evaluate how well the observed scores of the items are indicators for the latent (unobserved) constructs (Kieffer, 1999). To conduct a CFA, one must specify a theory-based model, select a method to estimate the model parameters, evaluate the goodness of fit of the theorized model via tests of the parameter estimates, and consider model modifications (Lei & Wu, 2007).

The theory-based model in this study is the SAS. One of the most popular estimation methods of CFA is maximum likelihood estimation (Lei & Wu, 2007). When this estimation procedure is used with an adequate sample size, appropriate model specification, and multivariate normality, it produces standard errors that allow researchers to assess the overall model fit (Schmitt, 2011). Unfortunately, maximum likelihood estimation is not appropriate for ordinal data due to the assumption of observed continuous variables and multivariate normality (Li, 2016). An alternative to maximum likelihood estimation is weighted least squares (WLS) estimation. The WLS estimation does not assume a specific distribution form; therefore, it can be applied to ordinal and continuous observed variables (Kline, 2011). One of the requirements of WLS is a large sample size; hence, Kline (2011) suggested the use of diagonal weighted least square (DWLS) estimation when the sample size is small. DWLS is specifically designed for ordinal data, taking into account that ordinal data are not continuous and non-normal (Li, 2016). This estimation is built on the framework of weighted least squares, but only incorporates the diagonal elements of the full weighted matrix (DiStefano & Morgan, 2014; Li, 2016). When conducting a DWLS estimation using LISREL, data can be inputted for estimation by way of the asymptotic variance or the full weight matrix (DiStefano & Morgan, 2014). To allow for the most unrestricted information, the full weight matrix was used to estimate parameters in this study.

In LISREL, the full weight matrix involves two steps: first, the estimation of the polychoric correlation matrix from the observed variables and second, the computation of the asymptotic variance matrix from the polychoric correlation matrix and other statistics (Kline, 2011; Mîndrilă, 2010). A polychoric correlation is a bivariate correlation of ordinal variables with three or more categories (Kline, 2011) and the asymptotic covariance matrix is a weight matrix containing the covariances of the observed sample variances (DiStefano & Morgan, 2014).

The goodness of overall model fit was determined via a chi-square test, root means square error of approximation (RMSEA), comparative fit index (CFI), and standardized root mean square residual (SRMR). The chi-square (χ^2) test is used to evaluate the difference between the sample covariance matrix and the model-implied covariance matrix. A non-statistically significant result indicates that there are no statistically significant differences between the two covariance matrices; the theoretical model statistically significantly reproduces the sample covariance matrix (Schumacker & Lomax, 2010). One of the assumptions of the χ^2 test statistic is multivariate normality (Kline, 2011). This assumption is violated when using ordinal variables. Hence, the Satorra-Bentler chi-square (χ^2_{SB}) to correct for non-normality is the preferred test for ordinal variables analyzed with a weighted least squares method of estimation (Kline, 2011). The RMSEA fit statistic informs how well the theorized model would fit the population's covariance matrix (Hooper, Coughlan, & Mullen, 2008). Hooper et al. (2008) suggested that RMSEA values between .10 and .08 indicate fair model fit, and values below .08 indicate good model fit. The CFI fit statistic is an incremental fit index that compares the sample covariance matrix with an independence model, a model with uncorrelated latent variables (Hooper et al., 2008; Kline, 2011). CFI values range from 0 to 1, with 0 representing no fit and 1 representing perfect fit. Although CFI values close to 1 indicate good model fit, Hu and Bentler (1999) suggest a minimum CFI value of .95 to ensure that misspecified theorized models are not accepted. The SRMR fit index is the difference between the sample covariance matrix and the model-implied covariance matrix; these values should be small (Kline, 2011). The suggested threshold for acceptable fit is a SRMR value less than .08 (Hu & Bentler, 1999). For this study, acceptable model fit is defined by a Satorra-Bentler chi-square corrected for a non-normality probability value greater than .05, RMSEA less than 0.08, CFI greater than .95, and SRMR less than .08.

Relationship to Other Variables

In addition to examining the structural validity, the convergent validity and discriminant validity of the SAS was assessed to ensure that (a) scores on items from the same factor are highly correlated with each other and (b) scores on items from different factors are not highly correlated to each other (Onwuegbuzie, Daniel, & Collins, 2009). To calculate the convergent validity, the square of the correlations (i.e., shared variance) between the scores on items of the SAS were examined. The convergent validity is the extent to which the shared variance between scores on items of the SAS that measure the same construct is more than the shared variance between scores on items that measure different constructs (Farrell, 2010). Likewise, the discriminant validity of the SAS was assessed to examine the extent to which scores from the factors of SAS are not correlated with students' scores from the factors of ATS. This type of validity was examined to ensure that the items on the SAS measure subscales exclusively related to statistics anxiety as designed (Vigil-Colet et al., 2008).

COMPREHENSIVE REFERENCE LIST

Baloglu, M. (2002). Psychometric properties of the statistics anxiety rating scale. *Psychological Report*, 90, 315-325. doi:10.2466/pr0.2002.90.1.315

Baloglu, M. (2003). Individual differences in statistics anxiety among college students. *Personality and Individual Differences*, *34*, 855-865. doi:10.1016/S01918869(02)00076-4

Baloglu, M., Deniz, M. E., & Kesici, S. (2011). A descriptive study of individual and cross-cultural differences in statistics anxiety. *Learning and Individual Differences*, *21*, 387-391. doi:10.1016/j.lindif.2011.03.003

Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology*, *87*, 611-623. doi:10.1037/0022-0663.87.4.611

Bell, J. A. (2001). Length of course and levels of statistics anxiety. *Education*, *121*, 713-716.

Bell, J. A. (2003). Statistics anxiety: The nontraditional student. *Education*, *24*, 157-162.

Bell, J. A. (2008). Statistics anxiety and business statistics: The international student. *Education*, *129*, 282-286.

Betz, N. E. (1978). Prevalence, distribution, and correlates of math anxiety in college students. *Journal of Counseling Psychology*, *25*, 441-448. doi:10.1037/0022-0167.25.5.441

Birenbaum, M., & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, *36*, 93-98. doi:10.1080/0013188940360110

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230-258. doi:10.1177/0049124192021002005

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington DC: American Psychological Association.

Cashin, S. E., & Elmore, P. B. (2005). The survey of attitudes toward statistics scale: A construct validity study. *Educational and Psychological Measurement*, *65*, 509-524. doi:10.1177/0013164404272488

Chew, P. K. H., & Dillon, D. B. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science*, *9*, 196-208. doi:10.1177/1745691613518077

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal*, *9*, 6-26.

Chiesi, F., Primi, C., & Carmona, J. (2011). Measuring statistics anxiety: Cross-country validity of the statistical anxiety scale (SAS). *Journal of Psychoeducational Assessment*, *29*, 559-569. doi:10.1177/0734282911404985

Chiou, C., Wang, Y., & Lee, L. (2014). Reducing statistics anxiety and enhancing statistics learning achievement: Effectiveness of a one-minute strategy. *Psychological Reports*, *115*, 297-310. doi:10.2466/11.04.PR0115c12z3

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319. doi:10.1037/1040-3590.7.3.309

Cruise, R. J., & Wilkins, E. M. (1980). *STARS: Statistical Anxiety Rating Scale*. Unpublished manuscript, Andrews University, Berrien Springs, MI.

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, Las Vegas, NV*.

DeVaney, T. A. (2010). Anxiety and attitude of graduate students in on-campus vs. online statistics courses. *Journal of Statistics Education*, *18*, 1-15.

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*, 425-438. doi:10.1080/10705511.2014.915373

Dunn, K. (2014). Why wait? The influence of academic self-regulation, intrinsic motivation, and statistics anxiety on procrastination in online statistics. *Innovative Higher Education*, *39*, 33-44. doi:10.1007/s10755-013-9256-1

Dykeman, B. F. (2011). Statistics anxiety: Antecedents and instructional interventions. *Education*, *132*, 441-446.

Earp, M. S. (2007). *Development and validation of the statistics anxiety measure* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses Global. (UMI No. 3279112).

Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty and Shiu (2009). *Journal of Business Research*, *63*, 324-327. doi:10.1016/j.jbusres.2009.05.003

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*, 466-491. doi:10.1037/1082-989X.9.4.466

Gardner, H. (1983). *Frames of mind*. New York, NY: Basic Book.

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 Individuals Differences*, *45*, 68-74. doi:10.1016/j.paid.2008.02.021

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha (Methods, plainly speaking). *Measurement and Evaluation in Counseling and Development, 34*, 177-189.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53-60.

Hsiao, T. (2010). The statistical anxiety rating scale: Further evidence for multidimensionality. *Psychological Reports*, *107*, 977-982. doi:10.2466/07.11.PR0.107.6.977-982

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equational Modeling*, *6*, 1-55. doi: 10.1080/10705519909540118

Jöreskog, K.G., & Sörbom, D. (1996). *LISREL user's reference guide*. Chicago, IL: Scientific Software International.

Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist*, *34*, 684-718. doi:10.1177/0011000006286347

Keeley, J., Zayac, R., & Correia, C. (2008). Curvilinear relationships between statistics anxiety and performance among undergraduate students: Evidence for optimal anxiety. *Statistics Education Research Journal*, *7*, 4-15.

Kesici, S., Baloglu, M., & Deniz, M. E. (2011). Self-regulated learning strategies in relation with statistics anxiety. *Learning and Individual Differences*, *21*, 472-477. doi:10.1016/j.lindif.2011.02.006

Kieffer, K. M. (1999). An introductory primer on the appropriate use of exploratory and confirmatory factor analysis. *Research in Schools*, *6*(2), 75-92.

Kline, R. B. (2011). *Principles and practice of structured equation modeling* (3rd ed.). New York, NY: The Guilford Press.

Lei, P., & Wu, Q. (2007). An NCME instructional module on introduction to structural equation modeling: issues and practical considerations. *Educational Measurement, Issues and Practice*, *26*(3), 33-43. doi:10.1111/j/1745-3992.2007.00099.x

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonal weighted least squares. *Behavior Research Methods*, *48*, 936-949. doi:10.3758/s13428-015-0619-7

Macher, D., Paechter, M., Papousek, I., Ruggeri, K., Fruedenthaler, H. H., & Arendasy, M. (2013). Statistics anxiety, state anxiety during an examination, and academic achievement. *British Journal of Educational Psychology*, *83*, 535-549. doi:10.1111/j.2044-8279.2012.02081.x

Macheski, G. E., Buhrmann, J., Lowney, K. S., & Bush, M. E. L. (2008). Overcoming student disengagement and anxiety in theory, methods, and statistics courses by building a community of learners. *Teaching Sociology*, *36*(1), 42-48. doi:10.1177/0092055X0803600106

Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, *1*(1), 60-66. doi:10.20533/ijds.2040.2570.2010.0010

Mji, A. (2009). Differences in university students' attitudes and anxiety about statistics. *Psychological Reports*, *104*, 737-744. doi:10.2466/PR0.104

Mji, A., & Onwuegbuzie, A. J. (2004). Evidence of score reliability and validity of the statistical anxiety rating scale among technikon students in South Africa. *Measurement and Evaluation in Counseling and Development*, *36*, 238-251.

Nasser, F. M. (2004). Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking pre-service teachers in introductory statistics. *Journal of Statistics Education*, 12(*1*). Retrieved from https://ww2.amstat.org/publications/jse/v12n1/nasser.html

Neumann, D. L., Hood, M., & Neumann, M. M. (2009). Statistics? You must be joking: The application and evaluation of humor when teaching statistics. *Journal of Statistics Education*, *17*(2), 1-16.

Nolan, M. M., Beran, T., & Hecker, K. G. (2012). Survey assessing students’ attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, *11*, 103-123.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Nunnally, J.C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Onwuegbuzie, A. J. (1997). Writing a research proposal: The role of library anxiety, statistics anxiety, and composition anxiety. *Library and Informational Science Research*, *19*(1), 5-33. doi:10.1016/S0740-8188(97)90003-7

Onwuegbuzie, A. J. (1998). Role of hope in predicting anxiety about statistics. *Psychological Reports*, *82*, 1315-1320. doi:10.2466/pr0.1998.82.3C.1315

Onwuegbuzie, A. J. (1999). Statistics anxiety among African American graduate students: An affective filter? *Journal of Black Psychology*, *25*, 189-209. doi:10.1177/0095798499025002004

Onwuegbuzie, A. J. (2000). Statistics anxiety and the role of self-perceptions. *The Journal of Educational Research*, *93*, 323-330. doi:10.1080/00220670009598724

Onwuegbuzie, A. J. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement*, *63*, 1020-1038. doi:10.1177/0013164402250989

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

Onwuegbuzie, A. J., Da Ros, D., & Ryan, J. M. (1997). The components of statistics anxiety: A phenomenological study. *Focus on Learning Problems in Mathematics*, *19*(4), 11-35.

Onwuegbuzie, A. J., & Daley, C. E. (1997, November). *The role of multiple intelligence in statistics anxiety*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Memphis, TN.

Onwuegbuzie, A. J., & Daley, C. E. (1999). Perfectionism and statistics anxiety. *Personality and Individual Differences*, *26*, 1089-1102. doi:10.1016/S0191-8869(98)00214-1

Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. T (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity*, *43*, 197-209. doi:10.1007/s11135-007-9112-4

Onwuegbuzie, A. J., Leech, N. L., Murtonen, M., & Tahtinen, 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

Papousek, I., Ruggeri, K., Macher, D., Paechter, M., Heene, M., Weiss, E. M., & Freudenthaler, H. H. (2012). Psychometric evaluation and experimental validation of the statistics anxiety rating scale. *Journal of Personality Assessment, 94*, 82-91. doi:10.1080/00223891.2011.627959

Pan, W., & Tang, M. (2004). Examining the effectiveness of innovative instructional methods on reducing statistics anxiety for graduate students in the social sciences. *Journal of Instructional Psychology*, *31*, 149-159.

Perepiczka, M., Chandler, N., & Becerra, M. (2011). Relationships between graduate students' statistics self-efficacy, statistics anxiety, attitude towards statistics, and social support. *The Professional Counselor: Research and Practice*, *1*, 99-108. doi:10.15241/mpa.1.2.99

Pretorius, T. B., & Norman, A. M. (1992). Psychometric data on the statistics anxiety scale for a sample of South African students. *Educational and Psychological Measurement*, *52*(4), 933-937. doi:10.1177/0013164492052004015

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, *74*, 525-556. doi:10.3102/00346543074004525

Robarte-Luna, B., & Sherry, A. (2008). Sex differences in the relation between statistics anxiety and cognitive/learning strategies. *Contemporary Educational Psychology*, *33*, 327-344. doi:10.1016/j.cedpsych.2007.03.002

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, *40*, 235-238. doi:10.1177/001316448004000138

Richardson, F. C., & Woolfork, R. (1980). Mathematics anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research, and applications* (pp. 271-288). Hillsdale, NJ: Erlbaum.

Russell, D. W. (2002). In search of underlying dimensions: the use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, 28, 1692-1646. doi:10.1177/014616702237645

Schacht, S., & Stewart, B. J. (1990) What's funny about statistics? A technique for reducing student anxiety. *Teaching Sociology*, *18*(1), 52-56. doi:10.2307/1318231

Schmitt, T.A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*, 304-321. doi:10.1177/0734282911406653

Schua, C. (2003, August). *Students attitudes: The other important outcome in statistics education*. Paper presented at the Joint Statistical Meetings, San Francisco.

Schau, C., Stevens, J., Dauphinee, T.L., & Del Vecchio, A. (1995). The development and validation of the survey attitudes toward statistics. *Educational and Psychological Measurement*, 55, 868-875. doi:10.1177/0013164495055005022

Schultz, K. S., & Koshino, H. (1998). Evidence of reliability and validity for Wise's attitude towards statistics scale. *Psychological Reports*, 82, 27-31. doi:10.2466/pr0.1998.82.1.27

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Routledge.

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, *19*, 80-110. doi:10.1177/1094428115598239

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. New York, NY: Routledge.

Teman, E. D. (2013). A Rasch analysis of the statistical anxiety rating scale. *Journal of Applied Measurement*, *14*, 414-434.

US College and University Directory (2016, August 11). *UnivSearch.com*. Retrieved from http://www.univsearch.com/family-search.php?family=45

VanHoof, S., Kuppens, S., Sotos, A. E. C., Verschaffel, L., & Onghena, P. (2011). Measuring statistics attitudes: Structure of the survey of attitudes towards statistics (SATS-36). *Statistics Education Research Journal*, *10*(1), 35-51.

Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L. (2008). Development and validation of the Statistical Anxiety Scale. *Psicothema*, *20*(1), 174-180.

Walsh, J. J., & Ugumba-Agwunobi, G. (2002). Individual differences in statistics anxiety: the roles perfectionism, procrastination and trait anxiety. *Personality and Individual Differences*, *33*, 239-251. doi:10.1016/S0191-8869(01)00148-9

Williams, A. S. (2010). Statistics anxiety and instructor immediacy. *Journal of Statistics Education*, *18*, 1-18.

Williams, A. S. (2013). Worry, intolerance of uncertainty, and statistics anxiety. *Statistics Education Research Journal*, *12*, 48-59.

Williams, A. S. (2014). An exploration of preference for numerical information in relation to math self-concept and statistics anxiety in a graduate statistics course. *Journal of Statistics Education*, *22*, 1-16.

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, *45*, 401-405.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73(6*), 913-934. doi:10.1177/0013164413495237

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, *61*, 319-328. doi:10.1111/j.2044-8279.1991.tb00989.x