CURVILINEAR RELATIONSHIPS BETWEEN STATISTICS ANXIETY AND PERFORMANCE AMONG UNDERGRADUATE STUDENTS: EVIDENCE FOR OPTIMAL ANXIETY

JARED KEELEY
Auburn University
keelejw@auburn.edu

RYAN ZAYAC
Auburn University
zayacrm@auburn.edu

CHRISTOPHER CORREIA
Auburn University
correcj@auburn.edu

ABSTRACT

This study examined the possibility of a curvilinear relationship between statistics anxiety and performance in a statistics course. Eighty-three undergraduate students enrolled in an introductory course completed measures of statistics anxiety and need for achievement at seven points during the semester in conjunction with six tests. Statistics anxiety scores were reliable internally and across time. Statistics anxiety decreased during the term yet paradoxically became more strongly related to performance. Curvilinear models were better predictors of test performance than linear, suggesting a mid-range optimal level of statistics anxiety. However, students' need for achievement proved not to mediate the relationship between anxiety and performance. The authors suggest ways these findings may influence future research in statistics anxiety and classroom management of anxiety.

Keywords: *Statistics education research; Statistics anxiety; Yerkes-Dodson law*

1. INTRODUCTION

Most students in the social sciences are required to take a statistics course as part of their program of study. However, anecdotally many of these students choose their particular majors in an attempt to avoid having to take "more math." As a result, students often dread their statistics course and may put it off until the end of their academic careers (Onwuegbuzie & Wilson, 2003; Roberts & Bilderback, 1980; Zeidner, 1991). Numerous authors have noted the presence of statistics anxiety among their students and its effects (Fitzgerald, Jurs, & Hudson, 1996; Onwuegbuzie & Seaman, 1995; Zanakis & Valenzi, 1997; Zeidner, 1991). There is a general consensus in the literature that statistics anxiety has an inverse relationship to performance in statistics classes (Fitzgerald et al., 1996; Onwuegbuzie & Seaman, 1995; Zanakis & Valenzi, 1997; Zeidner, 1991). For instance, Onwuegbuzie and Seaman (1995) found a negative correlation between statistics test

Statistics Education Research Journal, 7(1), 4-15, http://www.stat.auckland.ac.nz/serj

anxiety and students' final exam scores. Further, they found that there was an interactional effect with high anxiety students performing worse in timed conditions than in untimed conditions. This finding is consistent with research suggesting that the relationship between test anxiety and performance can be moderated by the complexity or difficulty of the exam, with high-anxious students performing best on easy or moderately difficult exams, and low-anxious students faring better on more difficult exams that enhance arousal and motivation (Zeidner, 1998).

The studies cited above have all examined linear relationships between anxiety and statistics performance. However, there is good theoretical reason to suggest that the relationship between anxiety and performance in the context of statistics may follow a curvilinear relationship. The well-known Yerkes-Dodson law (first described in Yerkes & Dodson, 1908) states that there is an optimal level of arousal for maximum performance. At both extremes of low and high levels of arousal, performance is poor. As arousal moves away from those extremes, performance gradually improves. Therefore, there is an optimal mid-range level of arousal. Thus, this relationship is curvilinear (more specifically, quadratic). The Yerkes-Dodson law has since been empirically validated in a variety of areas including trauma (McNally, 2003), sports performance (Kais & Raudsepp, 2004; Norton, Hope, & Weeks, 2004), stress on the job (Bhuian, Menguc, & Borsboom, 2005), artificial intelligence (Raudys & Justickis, 2003), animal research (Maes & de Groot, 2003), and most importantly for our purposes, in academic settings (Sarid, Anson, Yaari, & Margalith, 2004) and in relation to anxiety (Bodas & Ollendick, 2005; Hopko et al., 2003). Specifically, anxiety follows the same pattern as general arousal, in that low and high levels of anxiety are detrimental to performance in mental tasks (Hopko et al., 2003). In an academic setting, stress produced the same curvilinear relationship in performance as measured by students' grades (Sarid et al., 2004). Finally, anxiety seems to have the same effect on test performance (Bodas & Ollendick, 2005). Therefore, we expect that statistics anxiety will follow a curvilinear relationship with performance on statistics exams. This notion has been expressed before (Onwuegbuzie & Wilson, 2003) but has yet to be empirically tested.

Further, we expected that the possible curvilinear relationship between anxiety and performance may be moderated by other situational and dispositional factors. In the current study, we chose to focus on the potential effects of need for achievement. Research has generally shown that students with low academic motivation have lower grade-point averages (Cokley, Bernard, Cunningham, & Motoike, 2001; Vallerand et al., 1992). We hypothesized that a student's level of need for achievement (also known as achievement motivation) would moderate the relationship found between statistics anxiety and performance. We predicted that a student with a high level of achievement motivation would demonstrate the curvilinear relationship between anxiety and performance, whereas a student with a low level of achievement motivation would demonstrate no relationship. In the case of the highly motivated student, anxiety will be "fuel" for the student to perform, and so a moderate level of anxiety will produce the highest levels of performance on the test. However, we expect that students with a low need for achievement will be unaffected by their level of anxiety, as the anxiety will not be directed towards behaviors related to improving school performance (e.g., increased studying, asking for help, etc.).

The current study addressed three aims. First, the study examined the reliability of the Statistics Anxiety Ratings Scale (STARS) scores (Cruise & Wilkins, 1980), a commonly used measure of statistics anxiety (Onwuegbuzie & Wilson, 2003), with a sample of undergraduates taking an introductory level statistics course. Second, the study examined students' statistics anxiety across the term, specifically looking for a curvilinear relationship between anxiety levels and performance on statistics tests. Third, the study attempted to determine if students' level of achievement motivation was a moderating factor on the relationship between students' anxiety and performance.

2. METHOD

2.1. PARTICIPANTS

Participants were drawn from 83 students enrolled in a single introductory statistics course for the social sciences during the spring of 2005 at a large university located in the southeastern United States. The course was taught by one of the coauthors (CC), and the remaining coauthors (JK and RZ) served as the graduate teaching assistants. Students were required to take a basic level mathematics course as a prerequisite for enrollment in the course. Thus, the sample was one of convenience. Most students (73.5%) were female. The majority of students were seniors (71.1%), with some juniors (26.5%), two sophomores (2.4%), and no freshmen. Nineteen majors were represented, with the most frequent being psychology (24.1%), criminology (19.3%), and human development/ family studies (14.5%).

2.2. MEASURES

We administered two scales over the course of the study: the STARS (Cruise & Wilkins, 1980) and a modified version of the Work Value Survey's Achievement scale (Schwartz, 1994). The STARS consists of 51 items across six scales. The scales are designed to measure a student's (a) estimation of the worth of statistics (16 items), (b) anxiety regarding interpreting statistics (11 items), (c) test and class anxiety (8 items), (d) computational self-concept (7 items), (e) fear of asking for help (4 items), and (f) fear of the statistics teacher (5 items). Items are rated on two Likert scales ranging from 1 to 5 anchored as either "no anxiety" to "very much anxiety" or "strongly disagree" to "strongly agree." Higher scores on each scale are indicative of relatively higher levels of anxiety. Cruise, Cash, and Bolton (1985) reported internal reliability coefficients ranging from .68 to .94 for the subscale scores with re-test reliability ranging from .67 to .84. Of all the various measures of statistics anxiety that exist in the literature, the STARS is the most frequently used and most empirically investigated (Onwuegbuzie & Wilson, 2003).

We used the Achievement scale of the Work Value Survey (Schwartz, 1994) as a measure of students' need for achievement. The scale consists of six items rated on a 7 point Likert-format scale ranging from "opposed to my values" to "of supreme importance." Feather, Norman, and Worsley (1998) reported a reliability coefficient of .76 for the scale scores, and Schwartz (1994) presented some evidence of construct-related validity.

We also recorded students' performance on each of six non-cumulative tests across the semester. Each test consisted of 20 multiple-choice items and 2 to 4 open-ended problems requiring students to compute and interpret a statistical analysis. The multiple-choice portion of the test accounted for 60% of the students' test scores, and the remaining 40% was accounted for by their performance on the open-ended items. Each exam was worth 100 points, and the percentage of points earned on each exam was used in all analyses to standardize comparison across exams.

2.3. PROCEDURE

On the first day of class, students were introduced to the topic of statistics anxiety and informed that the experimenters (who were the teacher of record and the two TAs for the class) would be conducting a study on statistics anxiety throughout the course. The experimenters stated that students would be asked to complete a short questionnaire during the first lab meeting and after every test. In compensation, students would be given extra course credit for every time they participated. It was made clear that participation was optional and voluntary, that their decision to participate would in no way affect their status in the course, and that other opportunities for extra credit would be available over the course of the semester. To ensure confidentiality, students identified themselves on the questionnaires through use of a code name known only to them and the TAs. After all was explained, students were asked to sign a consent sheet indicating their permission for the experimenters to use their data as given on the surveys and their corresponding test scores. Therefore, there were seven administrations of the measures, once at the beginning of the course and directly after each of the six tests. At the first administration, 90% of students completed the surveys. Participation ranged from 82% to 76% for the following six administrations after every test. Due to missing data, differing numbers of students completed a particular measure at every assessment time.

3. RESULTS

3.1. STATISTICS ANXIETY RELIABILITY

The internal consistency of scores on the six scales of the STARS was good, with Cronbach’s alpha ranging from .83 to .94 (see Table 1). We examined the test-retest reliability of the scale scores during the term. Students’ level of statistics anxiety generally decreased during the term (see below for a discussion of this finding). We used standard Pearson correlations between scales at each time as the measure of reliability, as Pearson correlations are not affected by a score’s value but rather its relative position in relation to other scores at the same time. We assessed test-retest reliability in two ways. We examined the reliability between consecutive administrations, separated by approximately two weeks apiece, and we examined the reliability over the term between the first administration and the last. These values are presented in Table 1 for each of the scales. All the scale scores have good test-retest reliability across a two-week period, with average values around .8. The reliability coefficients drop when we consider the reliability over the term, but are still acceptably high given the time span of four months. All correlation coefficients were statistically significant at the .001 level.

Table 1. Internal and test-retest reliability coefficients for scores on the STARS scales

Scale	Cronbach’s α	Mean (Range) of correlations of consecutive pairs	Correlation of first and last administration
Worth of Statistics	.94	.82 (.71-.91)	.41
Interpretation Anxiety	.92	.87 (.83-.91)	.69
Test and Class Anxiety	.88	.84 (.79-.91)	.74
Computational Self-Concept	.88	.82 (.70-.90)	.64
Fear of Asking for Help	.83	.86 (.79-.92)	.69
Fear of Statistics Teachers	.85	.76 (.58-.89)	.61

3.2. STATISTICS ANXIETY AND TEST SCORES

Students' reported statistics anxiety followed a few interesting patterns over the course of the semester. First, scores on all of the scales uniformly dropped as the semester progressed. (See Figure 1.) A multivariate repeated measures ANOVA including each scale indicated that there were differences across the scales; Wilks' $\lambda(36, 955.68) = 0.49$, p-value < .001, partial $\eta^2 = .11$. The repeated aspect of the test required that only those students who completed the packet at all assessment times were analyzed ($n = 38$). The effect of attrition proved to be negligible as those students who completed the packet at every time differed from those who did not on only one measurement: the fear of the statistics teacher scale at the third assessment; $t(67) = -3.03$, p-value = .003, $\eta^2 = .12$. All other measurements at all times on all scales did not differ. The scores of each scale at each administration were adequately normal. However, the MANOVA did not evidence adequate sphericity, and so we examined the Greenhouse-Geisser correction for the univariate tests of each scale. These tests indicated that there was a statistically significant drop across time on each scale; worth of statistics $F(2.59, 95.71) = 14.59$, $\eta^2 = .28$; interpretation anxiety $F(3.91, 144.53) = 21.44$, $\eta^2 = .37$; test and class anxiety $F(3.93, 145.21) = 12.42$, $\eta^2 = .25$; computational self-concept $F(3.12, 115.57) = 7.86$, $\eta^2 = .18$; fear of asking for help $F(3.74, 138.31) = 11.74$, $\eta^2 = .24$; and fear of the statistics teacher $F(3.88, 143.48) = 3.02$, $\eta^2 = .08$; all p-values < .001 except for fear of the statistics teacher p-value = .021. Specific contrasts within the test above indicated that test and class anxiety scores were higher than all other scales; Fs(1, 37) range from 51.88 to 14.21, all p-values ≤ .001. The fear of the statistics teacher scale was lower than all other scales; Fs(1, 37) range from 51.88 to 10.38, p-values ≤ .003, except for the fear of asking for help scale; $F(1, 37) = 6.95$, p-value = .012 which was non-significant using a Bonferroni correction for the number of tests. All other scales were not statistically significantly different from each other. Table 2 presents the means and standard deviations for each scale across each administration.

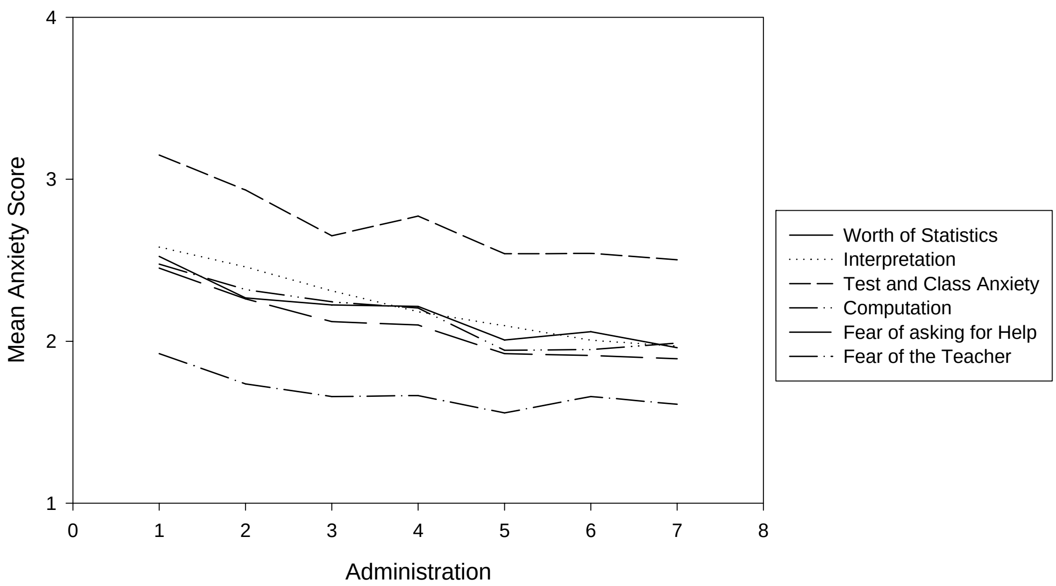

Figure 1. Students' average anxiety scores for each scale across the seven administrations

Table 2. Means (and standard deviations) for the STARS scales and the six tests

	Administration Time						
	1	2	3	4	5	6	7
Worth of Statistics	2.51	2.34	2.27	2.12	2.04	1.99	1.88
	(0.85)	(0.67)	(0.69)	(0.75)	(0.72)	(0.78)	(0.76)
Interpretation	2.46	2.37	2.18	2.00	1.95	1.94	1.79
	(0.83)	(0.74)	(0.72)	(0.75)	(0.77)	(0.77)	(0.68)
Test and Class Anxiety	3.12	2.84	2.60	2.68	2.51	2.47	2.45
	(0.92)	(0.82)	(0.97)	(0.94)	(0.90)	(1.00)	(1.04)
Computation	2.45	2.33	2.35	2.18	2.02	1.90	1.91
	(1.08)	(0.90)	(1.24)	(0.92)	(0.86)	(0.85)	(0.93)
Fear of Asking for Help	2.45	2.23	2.12	1.99	1.87	1.91	1.83
	(0.96)	(0.97)	(0.99)	(1.02)	(1.03)	(1.00)	(1.03)
Fear of the Teacher	1.91	1.82	1.88	1.65	1.66	1.68	1.64
	(0.84)	(0.71)	(0.80)	(0.80)	(0.76)	(0.91)	(0.89)
Test Scores		88.94	87.25	80.85	79.93	80.72	73.63
		(11.5)	(10.7)	(13.4)	(12.5)	(11.7)	(15.6)

A repeated measures ANOVA (n = 71) indicated that students' test scores also decreased across the term; $F(4.13, 289.38) = 29.31$, p-value < .001, $\eta^2 = .30$. Each test was normally distributed, but the test scores did not evidence adequate sphericity, and so we examined the Greenhouse-Geisser correction. Some test scores dropped more than others (see Figure 2). To examine these differential drops, we conducted post-hoc contrasts within the same repeated measures ANOVA. Students' performance on Test 1 was approximately equal to their performance on Test 2; $F(1, 70) = 1.77$, p-value = .19. However, there was a statistically significant decline from Test 2 to Test 3; $F(1, 70) = 31.38$, p-value < .001, $\eta^2 = .31$. Tests 3 and 4, as well as 4 and 5, were approximately equal; $F(1, 70) = .45$, p-value = .51 and $F(1, 70) = .33$, p-value = .57, respectively. However, there was a statistically significant drop from Test 5 to Test 6; $F(1, 70) = 23.82$, p-value < .001, $\eta^2 = .25$.

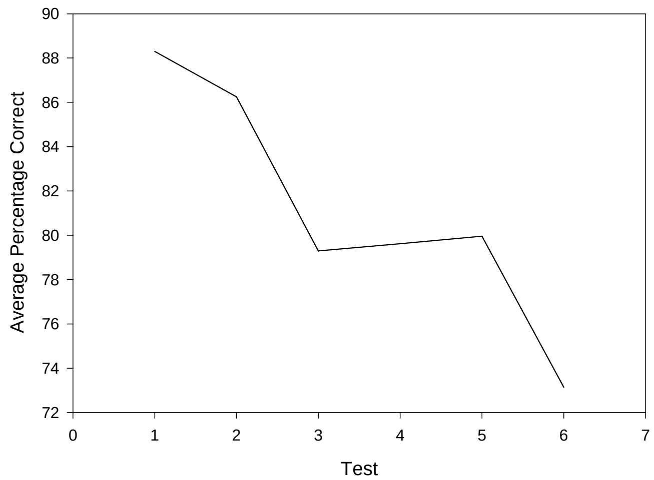

Figure 2. Students' average test scores across the six exams

Although it seems counterintuitive for students' anxiety to drop and for their corresponding test scores to also decrease, interestingly, students' anxiety scores became more related to their test scores as the term progressed and tests presumably became more difficult. At the time of the first test, none of the anxiety scales were statistically significantly correlated with test scores (*r* values ranging from -.080 to .009). At the second test, only computational self-concept was related to test scores; $r = -.26$, p-value = .03. On the second test, students were required to compute a standard deviation by hand, which may explain the relation of computational self-concept to performance. However, at Test 3, all anxiety scales except fear of the teacher were statistically significantly related to test scores, with lower anxiety being associated with higher test scores. The statistically significant correlations ranged from -.27 to -.47, all p-values < .05. The worth of statistics scale and computational self-concept scale were significantly related to the fourth test; *r* values = -.27 and -.35, p-values < .05, respectively. At Test 5, all scales except interpretation anxiety were related to test scores with *r* values varying between -.26 and -.46, p-values < .05. At Test 6 all scales significantly correlated with test scores; *r* values ranging from -.30 to -.45, p-values < .05. All of these relationships are moderate at best, but still represent notable effects. Interested parties may contact the authors for a copy of the full correlation matrix.

We hypothesized that there would be a curvilinear relationship between anxiety and test performance, with high anxiety and low anxiety being associated with low test scores and mid-level anxiety evidencing the best performance. Specifically, we expected the test anxiety scale to demonstrate most the non-linear relationship, as it is most directly related to test performance conceptually. To test this hypothesis, we conducted hierarchically nested regressions including only a linear term first (Model 1), followed by a model including both a linear and quadratic term (Model 2), so that we may examine incremental improvement in prediction. During the early tests, we saw that performance was not meaningfully related to anxiety, likely due to a ceiling effect with the scores on those two tests being almost a letter grade higher than the others. However, as the term progressed and exam scores dropped, anxiety and test performance became more strongly related. At Test 3, although the relationship between test anxiety and test performance was accounted for using a linear model (Model 1 $r^2 = .10$, $F(1, 61) = 6.74$, p-value = .01), a quadratic model demonstrated a statistically significant improvement (Model 2 $r^2 = .18$, $F(2, 60) = 6.44$, p-value = .003) with the individual quadratic term also demonstrating significance ($t(61) = -2.37$, p-value = .02). At Test 4, neither model was significant, but again the quadratic accounted for more variance (Model 2 $r^2 = .07$, $F(2, 61) = 2.37$, p-value = .10) than the linear (Model 1 $r^2 = .02$, $F(1, 62) = 1.59$, p-value = .21). For both Test 5 and Test 6, again the relationship with test anxiety was statistically significant using a linear relationship (for Test 5 Model 1 $r^2 = .07$, $F(1, 62) = 4.43$, p-value = .04; Test 6 Model 1 $r^2 = .09$, $F(1, 60) = 6.08$, p-value = .02), but the prediction was improved with the quadratic equation (for Test 5 Model 2 $r^2 = .16$, $F(2, 61) = 5.88$, p-value = .005 and for Test 6 Model 2 $r^2 = .15$, $F(2, 59) = 5.31$, p-value = .008). Both individual coefficients for the quadratic term were statistically significant (for Test 5 $t(62) = -2.63$, p-value = .01; for Test 6 $t(60) = -2.06$, p-value = .04). Further, two other anxiety scales were curvilinearly related to the sixth test. Worth of statistics demonstrated a statistically significant linear relationship (Model 1 $r^2 = .20$, $F(1, 60) = 14.95$, p-value < .001) and yet a quadratic term incrementally improved the prediction (Model 2 $r^2 = .25$, $F(2, 59) = 9.83$, p-value < .001); and interpretation anxiety followed the same pattern with the quadratic term improving upon the linear (Model 1 $r^2 = .11$, $F(1, 59) = 7.04$, p-value = .01; Model 2 $r^2 = .18$, $F(2, 58) = 6.49$, $p = .003$). It is possible that a quadratic equation could fit the data better than a linear and yet not follow the expected pattern (i.e., it could be shaped like a “u” rather than an “n”). All curvilinear relationships followed the pattern of extreme scores evidencing lower performance, with mid-level anxiety showing the highest performance.

As an example, the relationship between test anxiety and performance on Test 6 is depicted in Figure 3.

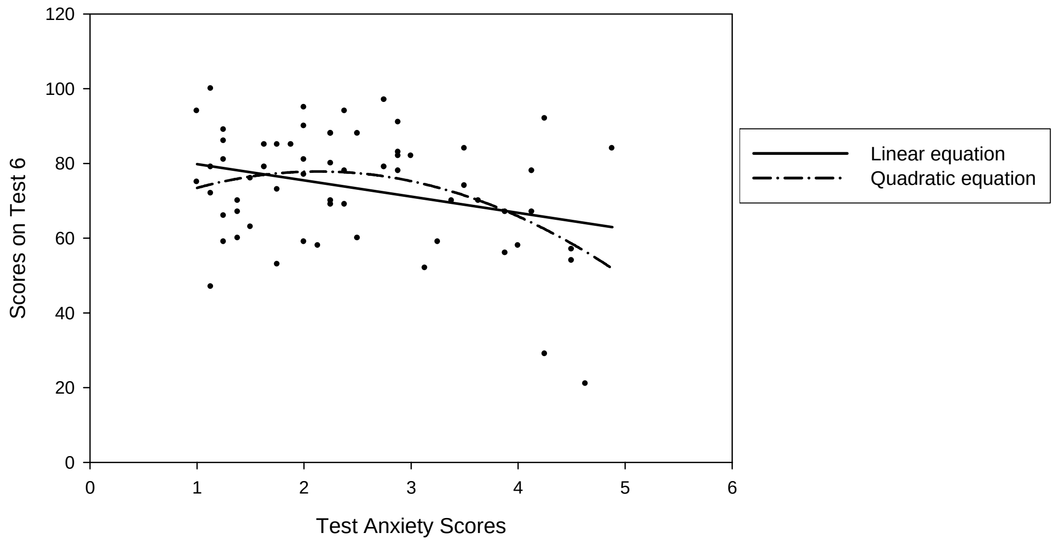

Figure 3. Linear and quadratic predictions of the relationship between test scores and anxiety at the seventh administration

3.3. NEED FOR ACHIEVEMENT

Students reported their need for achievement as starting at a moderately high level at the first assessment point (M = 5.50 out of a 7 point scale, SD = 0.84), but then dropped to a constant level for every assessment thereafter ($M \approx 4.5$, $SD \approx 1.0$). A repeated measures ANOVA using the 45 students who completed the measure at every assessment point indicated that this was a statistically significant decline; $F(6, 264) = 33.08$, p-value < .001, $\eta^2 = .43$.

However, this variable was only moderately related (all r values ≤ .29) to students' anxiety or test scores across the seven assessment times, with most correlations being true zero order relationships. Only a handful of these correlations were statistically significant (5 out of 48), and if a Bonferroni error correction is used to account for the number of tests conducted, they become statistically insignificant. In an attempt to determine whether differing levels of need for achievement affected test scores, we compared high versus low scorers on the scale as split by the mean score for that assessment time. Identical results occurred with a median split. The six t-tests were all statistically insignificant, with low and high achievement oriented students receiving nearly equal grades on their statistics tests. When low versus high achievement was entered into the repeated measures MANOVA of the six statistics anxiety scales as a between-subjects factor, it did not produce a main effect for any of the scales, nor was there a significant interaction across time for any scale.

4. DISCUSSION

A number of measures have been developed to assess statistics anxiety among undergraduates, including the Statistics Anxiety Scale (Pretorius & Norman, 1992), the Statistics Anxiety Inventory (Zeidner, 1991), and the STARS (Cruise & Wilkins, 1980).

In their review of available measures, Onwuegbuzie and Wilson (2003) noted that the STARS was the most extensively used and the only one subjected to concurrent validity testing (e.g., Baloglu, 2002; Onwuegbuzie, 2003; Walsh & Ugumba-Agwunobi, 2002), and that the reliability of scores on these measures had not been consistently reported. The current study was designed to evaluate the STARS further with a sample of undergraduates enrolled in a statistics course and addressed three aims. First, the study assessed reliability of STARS scores. Second, it examined the nature and strength of the relationship between statistics anxiety and performance over the course of a semester. Third, the study investigated the potential role of achievement motivation as a moderating factor on the relationship between statistics anxiety and performance.

As in previous studies (Baloglu, 2002; Cruise et al., 1985; Onwuegbuzie, 1998), all six of the original STARS scales displayed scores with good internal reliability. Our data also suggest that scores on each of the six STARS scales are reasonably reliable across seven administrations over the course of the semester, and our results are consistent with previously reported test-retest coefficients (Cruise et al., 1985). When taken as a whole, the existing literature supports the reliability and concurrent validity of the six commonly derived factor scores of the STARS.

A second aim of the study was to examine the relationship between statistics anxiety and performance over the course of the semester. Scores on all six of the statistics anxiety scales decreased statistically significantly over the course of the semester, suggesting that students became less anxious about the perceived value of learning statistics and their own abilities. Zanakis and Valenzi (1997) reported that business students enrolled in a second statistics course reported a decrease in anxiety related to understanding statistics and seeking help. However, the business students actually increased their reported lack of interest and devalued their perceived worth of statistics. They also reported some differences in end-of-semester anxiety ratings across four instructors with different teaching styles and philosophies. More research is needed to determine the degree to which various aspects of statistics anxiety change over time. Because neither our study nor the Zanakis and Velenzi study included a control group, it will be important for future research to determine if changes in anxiety levels are specifically due to enrollment in a statistics course and exposure to material. Other variables that might influence changes in anxiety levels, including prior experience, the structure of the course, the style of the instructor, and the career interests of the students, also should be considered systematically.

In addition to studying changes in anxiety levels over the semester, we assessed the relationship between statistics anxiety and performance over the course of the semester. Several studies have reported a negative relationship between statistics anxiety and course performance (Fitzgerald et al., 1996; Onwuegbuzie & Seaman, 1995; Zanakis & Valenzi, 1997; Zeidner, 1991). Onwuegbuzie and Wilson (2003) hypothesized that statistics anxiety may impair performance by interfering with students' ability to receive, concentrate on, and encode the terms and concepts presented in class. However, the authors also noted that a certain level of statistics anxiety may actually be beneficial if it motivates adequate preparation. Our results offer some support for this nonlinear relationship between statistics anxiety and performance. As hypothesized, a quadratic equation best captured the relationship between the test anxiety scale of the STARS and performance on the last four exams of the semester, with high and low statistics anxiety corresponding to lower test scores and mid-level anxiety corresponding to the best performance. The worth of statistics and interpretation anxiety scales also showed a curvilinear relationship with performance on the final exam of the semester. It is interesting that the relationship between test anxiety and performance became stronger and more curvilinear after the second exam, given that student performance was lower on all of the subsequent exams. Similarly, the sixth and final exam, which was also the exam on which students performed most poorly, occasioned a curvilinear relationship between performance and the worth of statistics and interpretation scales. Thus, more dimensions of statistics anxiety exhibited the curvilinear relationship with performance as the exams became increasingly difficult. As previously noted, similar patterns have been reported in the literature, with high-anxiety students performing worse under more stressful administration conditions and on more difficult tasks than their low-anxiety counterparts (Onwuegbuzie, 1995; Onwuegbuzie & Seaman, 1995; Zeidner, 1998). However, our data cannot address other factors that may have influenced the relationship between anxiety and performance. For example, as exams became more difficult over the course of the semester, it is likely that the demands from students' other academic courses also were increasing. Specifically, on the sixth exam, many students determined what score they needed to earn the grade they desired, and studied accordingly. Looking at extraneous factors that may moderate the relationship is particularly important given Onwuegbuzie's (2003) finding that academic course load is inversely related to statistics performance.

The literature has also investigated the relationship between statistics anxiety and a number of intrinsic variables, including achievement expectation, perfectionism, procrastination, trait anxiety, and state anxiety (Baloglu, 2002; Onwuegbuzie & Wilson, 2003; Walsh & Ugumba-Agwunobi, 2002). As a third and final aim, we were interested in determining what role, if any, achievement motivation might play in the relationship between statistics anxiety and performance. Contrary to our initial hypothesis, need for achievement was not reliably related to performance or statistics anxiety. Given the prevalence of statistics anxiety among student populations, research on both intrinsic and extrinsic factors that moderate the relationship between anxiety and actual performance appears warranted. Additional research on strategies for optimizing levels of anxiety, and managing the consequences of debilitating anxiety, is also clearly warranted.

A number of limitations are worth noting. First, although our findings are consistent with previous experimental research in noting that the relationship between anxiety and performance strengthened as exams became more difficult (Zeidner, 1998), our study did not manipulate or explicitly control for the difficulty of our exams. Second, we did not control for order effects when administering our packets. Third, we used a convenience sample composed primarily of Caucasian females enrolled in our university's College of Liberal Arts. Future research will need to determine the degree to which our findings generalize to more diverse samples of undergraduates. Finally, the current study was primarily descriptive and exploratory in nature. There are likely a myriad of factors, including task difficulty, student motivation, institutional environment, and others, that play a role in the relationship between anxiety and performance. Hopefully, future work will explore these relations.

Despite these limitations, the results of this study still pose interesting implications for the teaching of undergraduate statistics. Teachers may engage in a variety of techniques to manage their students' anxiety. For example, teachers may use humor or other gimmicks to reduce anxiety (Schacht & Stuart, 1990). The results of our study suggest that uniformly reducing students' anxiety may be detrimental. Anxiety is not a fire that needs to be stamped out for students to be successful in a statistics class. Some anxiety is acceptable. For students, simply knowing that some anxiety is acceptable and even helpful may stop them from catastrophizing and increasing the negative effects of the anxiety they do experience. It would be interesting for future research to address the effect of such an intervention in a statistics class.

To summarize, our results suggest that STARS scores are a reliable measure of statistics anxiety. Our results also suggest that the relationship between statistics anxiety and performance on in-class exams is quadratic, rather than linear, and the relationship between anxiety and performance becomes stronger as exams become more difficult. Finally and contrary to our initial hypothesis, achievement motivation did not moderate the relationship between statistics anxiety and performance.

ACKNOWLEDGEMENTS

Part of the research presented in this paper was published previously in a poster by the same authors entitled "Statistics anxiety: Curvilinear versus linear relationships" at the National Institute on the Teaching of Psychology, January 2006.

REFERENCES

Baloglu, M. (2002). Psychometric properties of the Statistics Anxiety Rating Scale. *Psychological Reports, 90*(11), 315-325.

Bhuian, S. N., Menguc, B., & Borsboom, R. (2005). Stressors and job outcomes in sales: A triphasic model versus a linear-quadratic-interactive model. *Journal of Business Research*, *58*(2), 141-150.

Bodas, J., & Ollendick, T. H. (2005). Test anxiety: A cross-cultural perspective. *Clinical Child and Family Psychology Review*, *8*(1), 65-88.

Cokley, K. O., Bernard, N., Cunningham, D., & Motoike, J. (2001). A psychometric investigation of the Academic Motivation Scale using a United States sample. *Measurement and Evaluation in Counseling and Development*, *34*(2), 109-119.

Cruise, R., & Wilkins, E. (1980). *STARS: Statistical Anxiety Rating Scale*. Unpublished manuscript, Andrews University, Michigan.

Cruise, R., Cash, R., & Bolton, D. (1985). Development and validation of an instrument to measure statistical anxiety. *ASA Proceedings of the Section on Statistical Education Section* (pp. 92-97). Alexandria, VA: American Statistical Association.

Feather, N. T., Norman, M. A., & Worsley, A. (1998). Values and valences: Variables related to the attractiveness and choice of food in different contexts. *Journal of Applied Social Psychology*, *28*(7), 639-656.

Fitzgerald, S., Jurs, S., & Hudson, L. (1996). A model predicting statistics achievement among graduate students. *College Student Journal*, *30*, 361-366.

Hopko, D. R., McNeil, D. W., Lejuez, C. W., Ashcraft, M. H., Eifert, G. H., & Riel, J. (2003). The effects of anxious responding on mental arithmetic and lexical decision task performance. *Journal of Anxiety Disorders*, *17*(6), 647-665.

Kais, K., & Raudsepp, L. (2004). Cognitive and somatic anxiety and self-confidence in athletic performance of beach volleyball. *Perceptual and Motor Skills*, *98*(2), 439-449.

Maes, J. H. R., & de Groot, G. (2003). Effects of noise on the performance of rats in an operant discrimination task. *Behavioural Processes*, *61*(1), 57-68.

McNally, R. J. (2003). Psychological mechanisms in acute response to trauma. *Biological Psychiatry*, *53*(9), 779-788.

Norton, P. J., Hope, D. A., & Weeks, J. W. (2004). The Physical Activity and Sport Anxiety Scale (PASAS): Scale development and psychometric analysis. *Anxiety Stress and Coping*, *17*(4), 363-382.

Onwuegbuzie, A. J. (1995). Statistics test anxiety and female students. *Psychology of Women Quarterly, 19*(3), 413-418.

Onwuegbuzie, A. J. (1998). Statistics anxiety: A function of learning styles? *Research in the Schools, 5*(1), 43-52.

Onwuegbuzie, A. J. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement, 63*(6), 1020-1038.

Onwuegbuzie, A. J., & Seaman, M. (1995). The effect of time constraints and statistics test anxiety on test performance in a statistics course. *Journal of Experimental Education*, *63*(2), 115-124.

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

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.

Raudys, S., & Justickis, V. (2003). Yerkes-Dodson law in agents' training. *Progress in Artificial Intelligence*, *29*, 54-58.

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, *40*(1), 235-238.

Sarid, O., Anson, O., Yaari, A., & Margalith, M. (2004). Academic stress, immunological reaction, and academic performance among students of nursing and physiotherapy. *Research in Nursing and Health*, *27*(5), 370-377.

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

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, *50*(4), 19-45.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, *52*(4), 1003-1017.

Walsh, J. J., & Ugumba-Agwunobi, G. (2002). Individual differences in statistics anxiety: The roles of perfectionism, procrastination, and trait anxiety. *Personality and Individual Differences, 33*(2), 239-251.

Yerkes, R., & Dodson, J. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurological Psychology*, *18*(3), 459-482.

Zanakis, S. H., & Valenzi, E. R. (1997). Student anxiety and attitudes in business statistics. *Journal of Education for Business*, *73*(1), 10-16.

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students – some interesting parallels. *British Journal of Educational Psychology*, *61*(3), 319-328.

Zeidner, M. (1998). Test anxiety: The state of the art. New York, NY: Plenum Press.

JARED KEELEY
Auburn University
Department of Psychology
226 Thach
Auburn, AL 36849, USA