The Stability of Undergraduate Students' Cognitive Test Anxiety Levels

Jerrell C. Cassady
Department of Educational Psychology
Ball State University

Test anxiety has been overwhelmingly identified as a two-factor construct, consisting of the cognitive (often referred to as "worry") and emotional (or affective) components (Morris, Davis, & Hutchings, 1981; Schwarzer, 1986). The predominant view of the relationship between these two factors suggests the cognitive component directly impacts performance (Bandalos, Yates, & Thorndike-Christ, 1995; Cassady & Johnson, in press; Hembree, 1988), while the emotional component is related but does not directly influence test performance (Sarason, 1986; Williams, 1991). The apparent relationship between emotionality and test performance is such that emotionality impacts test performance only under situations where the individual also maintains a high level of cognitive test anxiety (Deffenbacher, 1980; Hodapp, Glanzmann, & Laux, 1995). Although emotionality has traditionally not been viewed as central to performance, recent work has demonstrated that emotionality may be the triggering mechanism for self-regulation strategies that facilitate performance (Schutz & Davis, 2000).

This study investigated the stability of test anxiety over time by examining the level of reported cognitive test anxiety at three points in an academic semester (all proximally close to exams). The comparisons between levels of test anxiety over time were intended to track fluctuations in level of anxiety over time and across testing formats. The expectation was that cognitive test anxiety is a relatively stable (trait-like) construct, and that students' levels of anxiety would reflect a high degree of similarity within subjects over time. The underlying purpose of the study was to determine if it is indeed necessary to evaluate levels of test anxiety for each test taken, or if test anxiety is stable enough that evaluation at one point in time is sufficient for research that spans multiple exams. Because stability is largely influenced by the internal consistency or reliability of the measure in, it was also imperative to investigate the level of scale reliability at each point in the study. Under conditions where a scale is (a) internally consistent (reliable), and (b) demonstrates high levels of similarity in response over extended periods of time, it is reasonable to establish that the measured construct is stable (Nunnally & Bernstein, 1994).

Standard conceptualizations of the cognitive test anxiety construct have addressed the interplay between state and trait anxiety (Snow, Corno, & Jackson, 1996; Spielberger & Vagg, 1992). In this conceptualization, individuals with high levels of cognitive test anxiety generally hold heightened levels of trait anxiety, but in evaluative situations, their state anxiety also elevates (Zeidner, 1995). This combinatorial relationship can lead to feelings of anxiety that interfere with test performance through blocks to cue utilization, attenuated attentional resources, or mere cognitive interference from the worries and fears induced by anxiety (Geen, 1980; Hembree, 1988; Sarason, 1986). This relationship has also been characterized as an additive function of the dispositional and situational anxiety influences faced by students in evaluative scenarios (Zohar, 1998).

The attention to the situation-specific factors that lead to test anxious thoughts and behaviors has promoted the methodological practice of gathering test anxiety data as close to the testing event as possible, to best capture the contextual influences to anxiety in the research design (Cassady & Johnson, in press; Covington, 1985; Hodapp et al., 1995; Zeidner, 1998). Thus, the ideal time to test for test anxiety would be during an examination itself, with the subject providing "online" responses to the immediate feelings, fears, and behavioral responses that are arising during evaluation. However, this cannot occur in research designs that use students in actual testing situations, due to the probability of inducing additional debilitative cognitive test anxiety by having the student respond to items addressing their level of worry or fear for tests. Further, the implication is that in order to measure the performance-test anxiety relationship over a series of tests would require repeated administration of the test anxiety measures. These conditions are pragmatically undesirable, and potentially unnecessary.

Method

Participants

Undergraduate students in an introductory educational psychology course were the participants in this investigation. The participants completed the study instruments during one academic semester, as one option for completing course credit. Sixty-four undergraduate students participated in the three phases of study, with several participants completing...
only portions of the data. The participants were predominantly White (n = 62), with two Black students participating (which included all Black students available for participation). Consistent with the course population, there were 47 females and 17 males.

Materials

All measures in this study have been validated and found to have high levels of internal consistency (Cassady, 2001b; Cassady & Johnson, in press). However, there has been no investigation with the Cognitive Test Anxiety scale to identify the test-retest reliability, or the stability of test anxiety over time with this instrument.

The Cognitive Test Anxiety scale (Cassady & Johnson, in press) is a 27-item measure focused on only the cognitive domain of test anxiety, formerly referred to as worry. The cognitive domain includes the tendency to engage in task-irrelevant thinking during test taking and preparation periods, the tendency to draw comparisons to others during test taking and preparation periods, and the likelihood to have either intruding thoughts during exams and study sessions, or have relevant cues escape the learner’s attention during testing. In previous investigations of the Cognitive Test Anxiety scale, involving over 1,000 participants, a reliable method for determining high and low levels of test anxiety has been documented (Cassady, 2001a; 2001b; Cassady & Johnson, in press), which splits respondents into three levels of test anxiety. The low group scores range from 27 to 61, moderate test anxiety group ranges from 62 to 71, and responses that score over 72 are categorized as the high test anxiety group (maximum score possible = 108).

Sarason’s (1984) Bodily Symptoms subscale from the Reactions to Tests four-factor scale of test anxiety was used as a measure of the emotionality component of test anxiety. The Bodily Symptoms scale has been shown to have an acceptable degree of internal consistency, despite the short length (k = 10; Sarason, 1984; 1986).

Procedures

Throughout the course of one academic semester, students were invited to complete the Cognitive Test Anxiety scale and the Bodily Symptoms subscale three times. The timing of completion of these scales was aligned with the testing times in the course. All students completed the scales no more than seven days before they took the course examination. The time was variable due to the students’ ability to choose which day to complete the examination. All scales were completed in groups, in the students’ course classroom. Scale completion generally took between 8 and 15 minutes.

The first two course exams were 35-item multiple-choice tests, administered online at the students’ convenience. However, the third exam was a take-home, open book exam that had the same weight in overall course grade as the other two exams. Due to the methods of data collection that were designed to maintain student confidentiality as related to the test anxiety and emotionality responses, matching students test scores to the dependent variables in this study was not possible.

Results

Data analyses focused on examining the stability of the students' cognitive test anxiety reports, and their perceptions of emotionality as measured by bodily symptoms. Correlation analyses were used to illustrate the students’ stable reactions to the test events posed in the course. In addition, internal consistency for each scale was calculated to identify the level of reliability demonstrated in each administration of the two dependent measures. Finally, correction for attenuation of the correlations among the measures was conducted to provide an analysis of the hypothetical true score correlations given the condition that no measurement error were present in the study measures (Nunnally & Bernstein, 1994).

Initially, descriptive analyses on the test anxiety scores demonstrated that the students in this course had somewhat lower levels of test anxiety than previous uses of the scale (Cassady & Johnson, in press; Cassady 2001a; 2001b). The average cognitive test anxiety scores were similar for test 1 (M = 62.44, SD = 14.41, n = 59), test 2 (M = 62.46, SD = 15.39, n = 57), and test 3 (M = 61.49, SD = 14.94, n = 57), which all place the average score at the pre-established cut-off point between low and moderate cognitive test anxiety (Cassady, 2001a; 2001b). The mean scores on the bodily symptoms measure also were stable across the three administrations, and were somewhat lower than average. With a possible score range of 10 to 40, the average scores were in the low range for test 1 (M = 15.73, SD = 5.10, n = 56), test 2 (M = 16.41, SD = 5.73, n = 56), and test 3 (M = 15.48, SD = 5.62, n = 56).

To measure the stability of cognitive test anxiety and emotionality over the course of a semester, correlational analyses of the three points of administration were conducted (see Table 1). The results demonstrate that there is a very strong correlation between the students' reports of cognitive test anxiety across three points in the semester, as well as strong correlations among the three emotionality measurements. Further, the correlations between test anxiety and bodily symptoms are significant, and consistent with earlier research on the relationship between the two primary factors of test anxiety (Hembree, 1988). The correlations between cognitive test anxiety and bodily symptoms did reveal that the bodily symptoms scale administration that took place before the second exam was the strongest correlation for all three administrations of the Cognitive Test Anxiety scale. This deviates from the expectation that suggests contextual factors in place at each administration session would drive the cognitive and emotional components given together to be most similar. Although these correlational values do not vary greatly, the pattern may be due to the fact that the students
adjusted their reports of emotionality in response to the first course examination. This is supported by the fact that the second test administration period was also the point at which the Bodily Symptoms subscale mean was highest for the population.

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CTA (Test 1)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>.94 (59)</td>
<td>.96 (59)</td>
<td>.94 (59)</td>
<td>.63 (59)</td>
<td>.70 (59)</td>
<td>.51 (59)</td>
</tr>
<tr>
<td>2. CTA (Test 2)</td>
<td>.91 (52)</td>
<td>.95 (57)</td>
<td>.98 (57)</td>
<td>.56 (57)</td>
<td>.73 (57)</td>
<td>.54 (57)</td>
</tr>
<tr>
<td>3. CTA (Test 3)</td>
<td>.88 (53)</td>
<td>.93 (55)</td>
<td>.94 (55)</td>
<td>.52 (55)</td>
<td>.70 (55)</td>
<td>.58 (55)</td>
</tr>
<tr>
<td>4. BS (Test 1)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>.58 (56)</td>
<td>.52 (50)</td>
<td>.48 (50)</td>
<td>.91 (50)</td>
<td>.92 (50)</td>
<td>.87 (50)</td>
</tr>
<tr>
<td>5. BS (Test 2)</td>
<td>.64 (51)</td>
<td>.67 (56)</td>
<td>.64 (56)</td>
<td>.82 (56)</td>
<td>.88 (56)</td>
<td>.94 (56)</td>
</tr>
<tr>
<td>6. BS (Test 3)</td>
<td>.47 (52)</td>
<td>.50 (54)</td>
<td>.54 (54)</td>
<td>.79 (54)</td>
<td>.84 (54)</td>
<td>.91 (54)</td>
</tr>
</tbody>
</table>

Note: All p's < .001. Values on the diagonal indicate Cohen’s Alpha Inter-item reliability coefficient. Values on the top half of the table reflect the correlations after correction for attenuation, while the bottom half are not corrected. Values in parentheses report sample size for each analysis.

<sup>1</sup>Cognitive Test Anxiety scale score.

<sup>2</sup>Bodily Symptoms subscale score.

Finally, the internal consistencies of the two dependent measures were reported (see Table 1) for each of the three administration periods, using Cronbach’s Alpha (Nunnally & Bernstein, 1994). These measures of internal consistency confirmed previous work that demonstrated high levels of internal reliability, and were subsequently used to correct the correlational values for attenuation due to measurement error. The corrections for attenuation did not reveal any changes to the patterns of intercorrelations.

**Discussion**

The results demonstrate that it is methodologically practical to make use of test anxiety data gathered at times other than when a particular test in question is being completed, given that the data have been collected during a time in which typical test-induced contextual variables are activated. That is, it does not appear to be important to gather test anxiety data prior to every test upon which the anxiety construct is being used in research analyses, simply prior to any test. The high internal consistency paired with the high degree of correlation among repeated administrations of the Cognitive Test Anxiety scale and Bodily Symptoms subscale provide strong evidence that the constructs have long-range stability (Nunnally & Bernstein, 1994).

The data provide useful information regarding an efficient and methodologically sound approach for collecting test anxiety data from undergraduate students. It is reasonable to extrapolate from these results that test anxiety data collected in close proximity to an evaluative event can be used in analyses of the impact of test anxiety on any test within that academic period. The data also demonstrated that the Cognitive Test Anxiety scale, which has been shown to have high levels of internal consistency and high construct validity (Cassady & Johnson, in press), also provides stable and consistent measures of test anxiety over time and across testing formats.

One theoretical implication of these results also relates to the interpretation of test anxiety as a failure at multiple levels of information processing (Benjamin, McKeachie, Lin, & Holinger, 1981; Naveh-Benjamin, 1991; McKeachie, 1984). The research in this area has confirmed that students with high test anxiety are not only prone to failure in situations where time factors attenuate performance and retrieval of key information, but even in take-home examinations (Benjamin et al., 1981). Although exam performances were not available for this study, and no conclusions regarding the stability of a detrimental impact of test anxiety on performance can be drawn, the results demonstrate that the level of reported anxiety is consistent over time, despite variations in course exam format. That is, the level of anxiety induced by the
take-home examination did not differ significantly from the level of anxiety induced by the closed book multiple-choice examinations. Therefore, it seems that test anxious thoughts and behaviors are likely prompted by the presence of evaluative tasks, regardless of testing format. This finding extends previous work demonstrating no differential rates of cognitive test anxiety induced by in-class and online testing formats (Cassady, 2001a).

References


**Descriptors:** *Test Anxiety; Test Reliability; Test Construction; Test Validity*

**Citation:** Cassady, Jerrell C. (2001). The stability of undergraduate students' cognitive test anxiety levels. *Practical Assessment, Research & Evaluation, 7*(20). Available online: [http://PAREonline.net/getvn.asp?v=7&n=20](http://PAREonline.net/getvn.asp?v=7&n=20).