International Journal of Testing, 9: 3–19, 2009

ISSN: 1530-5058 print / 1532-7574 online
DOI: 10.1080/15305050902733448

Routledge
Taylor & Francis Group

Adapting the Cognitive Test Anxiety Scale for use with Argentinean University Students

Luis Alberto Furlan
Universidad Nacional de Córdoba

Jerrell C. Cassady
Ball State University

Edgardo Raul Pérez
Universidad Nacional de Córdoba

A new Spanish version of the Cognitive Test Anxiety Scale (CTAS) was created to be used explicitly with Argentinean university students. The scale was translated and verified through blind back translation and given to a large sample of students majoring in psychology or chemistry (N = 752). Exploratory Factor Analysis (N = 376) showed an internal structure of two factors that differed from the established English version of the CTAS. Examination of the items revealed that the factors were likely influenced by the phrasing of items that were originally designed to have several items require endorsement of low anxiety. Confirmatory factor analyses (N = 376) were conducted to compare the fit of three models for the scale. The results demonstrated that a 16-item single-factor solution was the preferable model. Further analyses demonstrated strong internal consistency, and test-retest stability of the short Spanish version. Results support the utility of the scale in future transcultural research on test anxiety with American and Argentinean learners.

Keywords: test anxiety, cross-cultural, Spanish, factor analysis

Cross-cultural research on test anxiety requires the use of measures that have been constructed with sufficient attention to the psychometric qualifications of

Correspondence should be addressed to Luis A. Furlan, Laboratorio de Evaluacion Psicologica y Educativa, Facultad de Psicologia, Universidad Nacional de Cordoba, Enrique Barros y Enfermera Gordillo, Ciudad Universitaria, Cordoba (5000) Argentina. E-mail: furlan@psych.unc.edu.ar.

reliability and validity in samples representing both native cultures. Attempting to interpret cultural differences and similarities with measures that have not been examined with respect to the equivalence in meaning across settings leads to erroneous conclusions and false attributions that mislead the research field. The purpose of this study was to lay the foundation for providing an empirical investigation of cognitive test anxiety with university students in Argentina, a population that has received little attention in multinational investigations. Specifically, this study validates the translation of the Cognitive Test Anxiety scale into Spanish, simultaneously revealing that items on the English version were inappropriate for this population. A revised 16-item version of the original form is provided for use in future research with this population.

Developing Measures for Cross-Cultural Test Anxiety Research

While there is a strong tradition of multinational studies in test anxiety, the corpus of research has been dominated by European and North American influences. Select bodies of research in Asia and the Middle East have been offered (Benson & El Zahar, 1994; Cassady, Mohammed, & Mathieu, 2004; Ikeda, Iwanaga, & Seiwa, 1996; Zohar, 1998), but virtually no published research has been provided that explores South American culture and test anxiety.

The first step in developing a quality cross-cultural investigation in test anxiety is to ensure that the measures employed in the new setting are appropriate for the target population. Frequently there are methodological biases driven by linguistic variances in meaning when measuring latent trait with self-report instruments (Van de Vijver & Tanzer, 1997). Specific to test anxiety, cross-cultural studies of test anxiety have demonstrated cultural differences in both the structure and the manifestation of test anxiety on related behaviors and perceptions (Benson, Moulin-Julian, Schwarzer, Seipp, & El-Zahhar, 1992; Cassady et al., 2004; Hocevar & El-Zahhar, 1992).

The most popular process for establishing a valid measure in a new culture is to translate an existing validated measure of test anxiety. The translation process requires particular attention to transliteral equivalence, which ensures that the translated version captures the intent of the original scale—not a mere literal translation of the words. To establish transliteral equivalence, fluent speakers in the original and new languages with personal experience in the target culture are optimal candidates for the translation process (see Benson et al., 1992; Cassady et al., 2004). First, one of the fluent bilingual speakers translates the original scale into the target language, using cultural conventions that are appropriate. After the initial version has been translated, a second bilingual speaker is given the translated version to provide a blind backtranslation. This backtranslation involves restating the items in the original language (with no prior exposure to the original scale to eliminate bias). Throughout the process, clarification with a third individual

regarding the intent of the original items is desirable; however, the two translators can have no direct contact during the translation process. After the new language version has been translated back into the original language, comparison of the new and original scales can identify items that require revision to maintain the meaning of the scale. Once the new language version is created, the next process involves standard psychometric analyses to revalidate and establish reliability for the new scale. To connect the new version to the demonstrated validity evidence existing for the original scale, the new language version must first be confirmed to have similar reliability and factor structure estimates. Alternatively, the process of examining the structure of a new scale can even lead to evidence for a new structure for examining anxiety in one or more cultures (Benson & El-Zahar, 1994).

Models for Measuring Test Anxiety

The international perspective of measuring test anxiety has been a long developing tradition, with a strong tradition of building on existing models and measures. In the beginning of empirical research in test anxiety, the construct was considered unidimensional and was assessed with scales such as the Test Anxiety Questionnaire (Mandler & Sarason, 1952). Subsequent research on test anxiety in students demonstrated that at least two dimensions of test anxiety were present. Liebert and Morris (1967) demonstrated that worry and emotionality in test anxiety were two distinct constructs. The worry component focuses on distracting thoughts, self-deprecating rumination, and other distractions to thought processes related to testing.

The emotionality component refers to biological and bodily responses that are related to anxiety (increased heart rate, headaches, perspiration; see Cassady, 2004a). The Test Anxiety Scale (Sarason, 1978) and Test Anxiety Inventory (Spielberger et al., 1980) were two popular assessment instruments developed in conjunction with this two-factor explanation for test anxiety. While the research confirmed the presence of the two factors repeatedly, unresolved psychometric problems persisted, including a strong overlap between the factors (Ware, Gallasi, & Dew, 1990; Ferrando, Varea, & Lorenzo, 1999). Another orientation toward dividing the test anxiety construct followed the standard state vs. trait research in anxiety.

Measuring anxiety in this tradition typically employs using general anxiety measures (State-Trait Anxiety Index; Spielberger, Gorsuch, & Lushene, 1970) in addition to test anxiety measures with the two factors represented (worry and emotionality). Sarason (1984) proposed that the standard two-factor model of anxiety was insufficient, recasting test anxiety measurement with the Reactions to Tests (RTT) scale, which includes four dimensions (10 items per dimension): tension, bodily symptoms, irrelevant thinking, and worry. Again, subsequent research was mixed, with some evidence for the unique impact of each aspect of the

RTT model (Flett, Blankstein, & Boase, 1987) and other data demonstrating the factorial representation was not durable or theoretically tenable (Rost & Schermer, 1992). Continued research on the RTT model produced shortened versions that appeared to support the model of a multidimensional construct for test anxiety in multinational investigations (Benson & Bandalos, 1992; Benson, Moulin-Julian, Schwarzer, Seipp, & El-Zahhar, 1992).

Cassady and Johnson (2002) proposed a measure of test anxiety (Cognitive Test Anxiety Scale) that addressed only cognitive indicators for test anxiety, encompassing the worry component of Liebert and Morris (1967) and aspects of the RTT model (Sarason, 1984). The contribution of the Cognitive Test Anxiety Scale (CTAS) to the field was deliberate attention to experiences prior to and following the testing event itself, consistent with process models of test anxiety that recognize the impact of anxiety across all phases of the learning-testing cycle (Cassady, 2004a). Research with the CTAS has acknowledged the presence of affective indicators of test anxiety through separate measures of emotionality (e.g., Sarason's Bodily Symptoms subscale from the RTT).

Previous research with the CTAS in Kuwait revealed the items could be translated into Arabic with only minor modifications (resulting in a 26-item Arabic version because of equivalence in meaning for two distinct items from the original scale, see Cassady et al., 2004). Other research focused on translating existing scales into new languages, which also required transformations to maintain the integrity of the assessment tool. Hoddap (1996) had to make several transformations to the original TAI scale, including new constructs, which resulted in four subscales: emotionality, worry, interference, and lack of confidence. The German Test Anxiety Inventory (G-TAI) includes 30 items describing experiences that occur only during the actual examination situation without temporal references to the participants' beliefs and behaviors prior to and following the actual testing session, in congruence with interference model. Replication studies with German samples validated the multifactor solution (Keith, Hoddap, Shermelleh-Engel, & Moosbrugger, 2003; Musch & Bröder, 1999), as well as providing evidence for a higher-order general test anxiety factor (Hoddap & Benson, 1997). In Spain, researchers have also adapted a translated version of Spielberger's Test Anxiety Inventory (Gutierrez, Calvo, & Avero, 1995; Gutierrez & Calvo, 1996) but there are not available data about its psychometrics properties. Valero Aguayo (1999) has built the CAEX (Test Anxiety Questionnaire) with 50 items related to four dimensions: behavioral, cognitive and physiological symptoms, and anxious situations. Ferrando,Varea, & Lorenzo (2003) tested the psychometric properties of Aguilar's (1984) CAR (Performance and Anxiety Questionnaire). Their analyses produced a three-factor model: worry, emotionality, and facilitating anxiety—a concept describing the motivational function of anxiety in evaluative tasks.

Present Investigation

While there is clearly a long history of adapting existing models and measures of test anxiety to explore test anxiety in a new culture, Furlan (2006) noted the absence of adapted measures for use in South American countries and articulated the need to develop effective scales for measuring anxiety in Argentina. While the naïve expectation would be that the Spanish versions created for use in Spain was a simple solution, linguistic and cultural differences may necessitate adaptations as laborious as translating English scale. Rather than create a "second-generation" translation and remove the Argentinean version further from the original measures, we determined translation of the original version was the prudent solution. To this end, the adaptation of CTAS to Spanish was undertaken with direct attention to creating a scale that was linguistically and culturally appropriate for Argentinean university students. Translation included the processes outlined earlier, including initial translation by native speakers from Argentina, verification of original scale meaning with the author of the original (English version) scale, blind back translation, and a psychometric investigation of the translated scale including an exploratory factor analysis, subsequent confirmatory factor analysis, and measures of reliability (internal consistency and stability).

METHOD

Following a comprehensive review of available test anxiety measures, the Cognitive Test Anxiety Scale (Cassady & Johnson, 2002) was selected by the Argentinean research team to be translated and used. The selection was made based on the available psychometric and theoretical evidence identifying the scale to be a reliable and valid measurement tool for examining cognitive test anxiety across the learning-testing cycle, accounting for test anxiety having an impact in the students' beliefs and behaviors during the test preparation phase, test performance phase, and test reflection phases (Cassady, 2004a; Schutz & Davis, 2000).

In addition to the psychometric adequacy of the scale, the singular focus on the cognitive aspects (e.g., worry) of test anxiety provide a more direct analysis of the aspect of test anxiety repeatedly shown to have the most durable and meaningful impact on actual academic performance (Cassady, 2004b; Morris, Davis, & Hutchings, 1981).

Cognitive Test Anxiety Scale

The Cognitive Test Anxiety Scale (CTAS) was created explicitly to focus on the cognitive aspects of test anxiety given repeated evidence that cognitive aspects of test anxiety were the most pervasive and reliable components affecting student

performance (e.g., Benjamin et al., 1981; Everson, Smodlaka, & Tobias, 1995; Naveh-Benjamin et al., 1987). Initial creation of the CTAS involved reducing a 44-item inventory that addresses cognitive test anxiety and identifying a stable unitary measure of cognitive test anxiety with studies with over 400 university students. The factor analyses, item analyses, and expert review of the CTAS resulted in the current 27-item version.

The focus of the CTAS is on 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 examinations and study sessions or to have relevant cues escape the learners' attention during testing. Measures of internal consistency repeatedly validate the reliability of the CTAS, with published values exceeding .91 (Cassady & Johnson, 2002; Cassady, 2004a). In addition, in a study of test-retest reliability, the CTAS was shown to be stable across three administration points in one academic semester (r's ranged from .88–.94; Cassady, 2001a). Responses to the CTAS range on a four-point Likert-type scale from "Not at all typical of me" to "Very typical of me." The range of possible scores in the original English version is 27 to 108. Because the CTAS contains items indicating both high and low levels of test anxious behavior, select items are recoded so that high scores indicated high levels of cognitive test anxiety.

Spanish Adaptation of CTAS

Following the standard translation procedures outlined earlier, the 27 items of the CTAS were translated into Spanish by a native speaker from Argentina, with explicit attention to transliteral equivalence. As such, periodic clarifications between the translator and the author of the English version were established. The nearest Spanish equivalent for casual English expressions in the original CTAS (e.g., "relaxed" or "I find myself thinking or worrying") were adopted because literal translation of such items would have produced expressions that would hold little relevance to native Argentinean speakers. Similarly, references to particularly specific conditions of study in universities in the United States (e.g., "assigned chapter in text books") were replaced by the Argentinean equivalents. After the Argentinean version was completed, the author of the English version recruited a native Argentinean teaching modern languages at a U.S. university to provide the backtranslation from Spanish to English. The second translator had no access to the original English version, producing a blind backtranslation condition. The original English version and the backtranslated version were compared for coherence by the English version author. Two items had inexact meaning in the translation and were reconciled among the translating team. In a pilot study, the newly translated Spanish CTAS (S-CTAS) was administrated to a small sample ($n = 10$) of Argentinean university students to verify the coherence of the scale instructions

and procedures as well as to seek feedback on the meaning of the items from the target population. The students in the Argentine pilot study identified that items requesting that they compare themselves with "the average college student" did not translate to their experiences well. That is, they reported that social referenced comparisons were atypical ways of thinking about their own abilities. Items with these self-other ability comparisons were retained in the Spanish version, partly because of the recognition that the tendency to avoid thinking in comparative terms may be a cultural factor worthy of comparison.

Validation Sample

The final step in this study was to examine the factor structure of the S-CTAS to identify similarity and difference with the English version and to explore the construct validity of the S-CTAS. To accomplish this, a large sample of Argentinean university students (n = 752) completed the Spanish version as a paper-pencil, anonymous, self-report task. Students completed the scale in group settings in 8 to 15 minutes on average. No incentives were provided for participation, and all participants were volunteers. Sixty percent of the sample were students with a major focus area in chemistry, with the remaining 40% majoring in psychology. The sample was split into two random halves. The first half of the sample was used to conduct an exploratory factor analysis, followed by a confirmatory factor analysis with the remaining sample. Further psychometric analyses confirmed construct validity and scale reliability. Specifically, the following common test anxiety patterns were investigated: stability over time, gender differences, and correlation with student academic performance.

RESULTS

Exploratory Factor Analysis

An Exploratory Factor Analysis (EFA) was run with 26 of the 27 items for the Spanish version (see Table 1). One item was removed from the analysis because the distribution of student responses on the item demonstrated violation of the assumption for normality, making inclusion in the EFA an inappropriate procedure. The item addresses the student's feeling that difficult items spur a sense of challenge rather than foreboding threat, and generally was shown to be consistent with low-test anxiety. However, conceptually, the item does have a risk of measuring a different construct than cognitive test anxiety given that it targets test confidence in addition to the absence of test anxiousness. This item was excluded in all analyses from the Spanish version. Values of the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO = .880) and Bartlett's test of sphericity ($X2 = 3055.82$, $df = 325$,

TABLE 1
EFA Pattern Matrix for 26 items Spanish CTAS

		N = 376	
Item #	Abridged Item Content	Factor 1	Factor 2
1	Lose sleep worrying about exams	.306[a]	
2	During exam, think others are doing better	.309[a]	
3	**Less difficulty than average students with instructions**		.393
4	Tend to freeze on intelligence tests or final exams	.481	
5	**Less nervous about tests than average student**		.712
6	During test, think about consequences of failing	.543	
7	At beginning of test, so nervous can't think straight	.560	
8	**Thought of taking test in a course would not cause worry**		.364
9	**More calm in tests than most**		.783
10	**Less difficulty learning textbook assignments**		.446
11	Mind goes blank when pressured for answer on test	.498	
12	During test, frequently think I'm not bright	.654	
13	**Do well in speed tests with time limits**		.272[a]
14	During test, I forget what I really know	.651	
15	After tests, I believe I could have done better	.652	
16	Worry more about doing well on tests than I should	.427	
17	**Before test, I feel confident and relaxed**		.686
18	**During test, I feel confident and relaxed**		.550
19	During test, have feeling I'm not doing well	.628	
20	When taking difficult test, feel defeated before I start	.550	
21[b]	**Finding unexpected items causes me to feel challenged rather than panicky**	—	—
22	I'm a poor test taker: performance does not show what I really know	.552	
23	I am not good at taking tests	.661	
24	When I get a test, it takes a while for me to calm down	.359	
25	I feel under pressure to do well on tests	.364	
26	I do not perform well on tests	.503	
27	When I take a test, nervousness causes me to make errors	.638	
	Rotated Eigenvalue	5.873	4.161
	Coefficient alpha	.88	.75

Note: Maximum likelihood, Promax Rotation employed. Bolded items are the 9 "reverse coded" items that assess low anxiety responses. **a:** Item factorial correlation < .30; **b**: Item excluded from EFA because it did not meet the assumption of normal distribution.

$p < .0001$) met necessary assumptions for conducting the EFA using maximum likelihood method.

Given the wealth of existing research in test anxiety that there are large correlations among various anxiety factors, we chose the oblique Promax rotation. The initial EFA solution contained six factors with eigenvalues exceeding the standard

1.0 value. The solution explained 55.29% of the variance. Examination of the scree plot suggested that the final four factors were not meaningful contributors to the factorial structure. The two dominant factors in the EFA accounted for 35.83% of the variance. More importantly, the two factors were coherent and easily defined. The first factor included 18 items (eigenvalue 6.23); the second included 7 items (eigenvalue 1.79). As anticipated, the two factors were moderately correlated ($r = .51$). One item was dropped from the solution because its factor loadings were below the criterion value of .30. Internal consistency estimates of the first and second factors revealed both scales were reliable, with coefficient alpha values of .88 and .75 (respectively).

Examination of the items revealed that the factor structure generated in the EFA was driven by syntactic differences among the items. That is, the original English-version CTAS was developed to include "reverse-coded" items in an attempt to avoid response bias. Several items were written such that a response of "Very Typical of Me" would be an indication of low anxiety, rather than the high anxiety indicators of the bulk of the items. Careful review of the items in the second factor revealed that the second factor was composed entirely of those reverse-coded items. This pattern raises the possibility that the factor structure is biased by either response confusion or that the reverse-coded items actually measure the "test confidence" rather than low-test anxiety. Other research on scale creation and development has pointed to similar "methods effects" in recent years (DiStefano & Motl, 2006), which reveal that the format of the items unduly influences the responses of the subjects.

Given the discovery of potential methods effects caused by the reverse-coded items, the EFA was re-run with only the 18 "high anxiety" phrased items. The scree plot suggested that one factor was the most reasonable solution (eigenvalue = 5.49) The factor solution revealed that only 17 of the 18 items were reasonably included in the final solution, which explained 29.88% of the variance. The item that fell out of the factor solution was one of the items requiring self-other comparisons, which the pilot sample identified as an atypical manner of thinking about test performance or ability.

Examination of the 2 EFAs revealed that one other item repeatedly produced the lowest factor loadings and just barely met the least stringent criteria for inclusion. This item (number 25) targets the student's perception of "feeling under pressure to do well" on examinations. In addition, reliability analyses of the scale revealed that the internal consistency of the scale is higher with the removal of that item. As such, the factor solution most supported by the data was a 16-item version including only the positively phrased items (see Table 2). Finally, a correlation between students' scores on the full-scale (27-item) and 16-item version of the S-CTAS was high and significant ($r = .94$, $p < .0001$), showing that shortened version could be considered an equivalent measure, with a more clear internal structure.

TABLE 2
EFA Factor Loadings for
16-Items CTAS

Item	Factor 1
1	.443
2[a]	.301
4	.559
6	.576
7	.725
11	.563
12	.564
14	.747
15	.543
16	.444
19	.592
20	.572
22	.388
23	.599
24	.429
25[b]	.360
26	.457
27	.727
Eigenvalue	5.977
Coefficient Alpha	.88

Note: a and b: Items excluded from final model.

Confirmatory Factor Analysis

An additional 376 students majoring in psychology or chemistry completed the 27-item version of the S-CTAS for the purposes of the confirmatory factor analysis. These students completed the survey at the same time as the EFA sample, but they were not included in those analyses.

The purpose of CFA was to estimate the fit of three alternative models generated in the EFA. The three models in review were (a) the original 26-item single factor scale, (b) a two-factor scale including both cognitive test anxiety and test confidence, and (c) a one-factor 16-item cognitive test anxiety scale.

To conduct the CFA, the AMOS statistical package was used, employing maximum likelihood analysis. As expected based on the EFA, the single-factor 26-item model representing the original structure of the CTAS did not have acceptable indices of fit (see Table 3). Overall, the indicators for goodness of fit for the data demonstrated minimal differences between the two alternative models. While the values for the CMIN/DF and RMSEA showed a slight advantage in the two-factor

TABLE 3
Fit Indices for Competing Models of the Spanish CTAS

MODEL	1 Factor (26 items)	2 Factor (26 items)	1 Factor (16 items)
CMIN/DF	4.059	3.127	3.461
CMIN	1213	931	359
DF	299	298	104
GFI	0.772	0.830	0.894
CFI	0.694	0.788	0.865
RMSEA	0.091	0,076	0.082
AIC	1317.583	1037.817	423.935
BIC	1690.507	1417.911	637.890

model, the GFI, CFI, AIC, and BIC were better in the 16-item model. Given the different lengths of the scales in review for these two CFA solutions, the AIC and BIC indicators hold merit in making final judgment. Furthermore, comparing the standardized coefficients from the factor analyses reveals that in only the 16-item model do the values all exceed the value .40 (see Table 4)

Taking into consideration the minimal differences in the models, the acceptable fit for the 16-item scale (Hu & Bentler, 1995) and the interest in maintaining parsimony with the construct of cognitive test anxiety, the selection of the 16-item version is most desirable for future uses with the S-CTAS.

Stability Estimates of Spanish CTAS

The original English version of the CTAS (all 27 items) was reported to be strong in three test administrations across one academic semester, with *r*s ranging from .88 to .95 (Cassady, 2001a). To confirm that the Spanish 16-item version was also stable over time, a separate test-retest reliability study was conducted. Fifty-nine Argentinean university students in psychology courses volunteered to complete the scale two times with a 90-day interval between administrations. The Pearson correlation coefficient revealed acceptable test-retest reliability ($r = .768$, $p < .01$). Comparisons between the two test administrations revealed similarity in group responses for the first ($M = 32.86$, $sd = 9.63$) and second administration ($M = 32.97$, $sd = 8.91$).

Gender Differences

Previous work with test anxiety and the CTAS in particular have routinely identified that females hold higher levels of anxiety over tests (Cassady & Johnson, 2002; Hembree, 1988). To examine gender differences on the reported levels of cognitive test anxiety in the Argentine sample, data from 371 students (71.7%

TABLE 4
Standardized Coefficients for the Three Alternative Models of Spanish CTAS

Item.	26 items 1 factor Factor 1	26 items 2 factors		16 items 1 factor Factor 1
		Factor 1	Factor 2	
1	.54	.53		.52
2	.31	.32		—
3	.22		.36	—
4	.60	.61		.61
5	.38		.50	—
6	.50	.50		.48
7	.76	.75		.74
8	.27		.40	—
9	.49		.70	—
10	.22		.33	—
11	.66	.68		.69
12	.58	.60		.60
13	.35		.28	—
14	.73	.74		.74
15	.50	.52		.52
16	.49	.49		.47
17	.50		.76	—
18	.52		.75	—
19	.55	.45		.56
20	.55	.56		.56
22	.35	.39		.42
23	.59	.60		.60
24	.48	.49		.49
25	.28	.30		—
26	.47	.49		.50
27	.63	.64		.63

female) were analyzed. The results of the independent sample t-test revealed a statistically significant difference, $t(269) = 2.49$, $p < .013$. The small, but significant difference revealed that females ($M = 35.25$, $sd = 9.14$) reported higher levels of cognitive test anxiety than the males ($M = 32.60$, $sd = 9.40$) on the 16-item Spanish version.

Criterion Validity: Test Performance

High levels of test anxiety have been associated with reduced academic performance; indeed, the primary reason for attention to test anxiety is the detrimental impact on test performance (Cassady & Johnson, 2002; Gutierrez & Calvo, 1996; Hembree, 1988). To examine the criterion validity of the Spanish CTAS, we

conducted an investigation of the relationship between student self-reported grade performance and test anxiety. To test the correlation between performance and test anxiety, 427 Argentinean university students studying chemistry completed the Spanish CTAS and provided a self-reported value for academic grade point average (GPA). Previous research with similar methodology has demonstrated the reliability and validity of using self-reported GPA as an accurate measure of university student performance (Cassady, 2001b). The results confirmed the criterion validity of the Spanish CTAS with a statistically significant, moderate, and negative correlation, $r = -.37$, $p < .01$.

DISCUSSION

The results of this study confirmed that the Spanish CTAS was a reliable and valid measure of cognitive test anxiety in the Argentinean university sample. The adaptation process and validation procedures provided not only evidence to support using the scale but also revealed a new factor structure that promotes the use of a shorter, superior form. Our results confirmed that the 16-item version of the Spanish CTAS was the most viable measure for the cognitive test anxiety construct.

Factorial Representation of Cognitive Test Anxiety

This single-factor solution is consistent with Cassady's (2002, 2004a) original conceptualization of cognitive test anxiety developed through work with the English version of the CTAS. The model did not provide support for breaking the cognitive test anxiety construct into subcomponents such as interference, lack of confidence, worry, or cognitive processing failures as suggested in other recent models (e.g., Hoddap, 1996; Sarason, 1984). The CTAS includes items that measure these various proposed separate factors, but our data maintain a strong fit with the single-factor model. Naturally, further research comparing the single-factor and multi-factor models directly is warranted to establish that the findings in this study apply with the other language versions of the CTAS. Adaptation to Spanish of G-TAI and RTAS could be necessary to cross-validate the findings, making studies in multi-national samples to test different models of the test anxiety construct.

The 16-item Spanish Cognitive Test Anxiety Scale (S-CTAS) was a shortened but equivalent version of the full scale with removal of two non-normally distributed items and all the original reverse-coded items. The two items that did not function appropriately in the Spanish version were likely not operating in the same way as in the original English version due to the social-referencing effect noted in the pilot study, in which Argentine students indicated difficulty conceptualizing

their own concerns over tests as a function of fellow classmates' abilities or performances. Previous research with the Cognitive Test Anxiety scale translated into Arabic and delivered to females from Kuwait revealed a similar trend—the pervasive social-self referencing noted in U.S. samples were not replicated in a collectivist society (Cassady et al., 2004). The reverse-coded items in the original CTAS were found to fall aside in a series of factor analyses. This outcome is consistent with the effects overviewed by DiStefano and Motl (2006), who demonstrated the presence of methods effects in survey research. In essence, the most evident explanation for the pattern of effects is that the respondents develop a response pattern to the scale, and reverse-coded items represent test anxiety in a different fashion. Close examination of the reverse-coded items suggests that they measure test confidence rather than serving as counter-indicators of the presence of test anxiety. Regardless, given the strong correlations among the shortened and full scale in both the Spanish and English versions, the shortened form is preferable as a more methodologically pure and parsimonious measure of cognitive test anxiety.

Establishing Transliteral Equivalence

The bulk of work in this study was in the process of translating an existing and validated measure into a new language, with specific attention to meeting the cultural norms of a country with little existing research on test anxiety. Our results confirmed prior findings for the need to tailor items to the language and sensibilities of the target population through careful translation procedures. In addition to the established procedures for scale translation, our experience repeatedly validated the value of communication and clarification between the scale authors. That is, clarifying the intent of the original English version with the author of that scale facilitated the development of a truly equivalent measure during the translation process. This collaborative process also enabled the discovery of the shortened version of the CTAS (in both languages), a finding that normally would have been overlooked or assumed to be a cultural effect rather than a persistent methodological effect.

CONCLUSION

In conclusion, this study confirms that the short version of the Spanish CTAS provides a valid and reliable measure of student cognitive test anxiety in a university setting. The establishment of a consistent and equivalent Spanish-language version of the CTAS provides a viable tool for conducting valid multinational and cross-cultural research on test anxiety, an area of research that is currently being pursued with these scales. The development and refinement of the Spanish

version was a necessary procedure in order to verify that those differences noted in forthcoming cross-cultural studies were based on actual differences between the target populations, and were not mere differences driven by measurement factors.

In addition to creating and validating the S-CTAS, the results of this study demonstrated that the methods effect was accounting for an improper fit in the factor solution for the CTAS. This discovery has prompted re-examination of the English language version of the CTAS, with preliminary analyses demonstrating that shortened version of the English version is also a preferred solution for measuring cognitive test anxiety. In addition, the data for the English version are also converging on the results from this study regarding items that drop out due to low factor loadings or non-normality in distribution. Further work on validating the CTAS is currently ongoing as we establish the optimal shortened version measure of cognitive test anxiety in both English- and Spanish-speaking samples.

REFERENCES

Aguilar, A. (1984). Cuestionario de Ansiedad y Rendimiento. Doc. Mimeog de uso exclusivo para la investigación. Tarragona: Seminario de Psicología.

Benjamin, M., McKeachie, W. J., Lin, Y., & Holinger, D. P. (1981). Test anxiety: Deficits in information processing. *Journal of Educational Psychology, 73,* 816–824.

Benson, J. & Bandalos, D. (1992). Second-Order Confirmatory Factor Analysis of the Reactios To Test Scale with Cross Validation. *Multivariate Behavioural Research, 27* (3), 459–487.

Benson, J., & El Zahar, N. (1994). Further refinement and validation of the Revised Test Anxiety Scale. *Structural Equation Modeling, 1,* 203–221.

Benson, J., Moulin-Julian, M., Schwarzer, C., Seipp, B., & El-Zahhar, N. (1992). Cross validation of a revised test anxiety scale using multi-national samples. In K. A. Hagtvet & T. B. Johnsen (Eds.) *Advances in test anxiety research* (Vol. 7, pp. 62–83). Lisse, The Netherlands: Swetts & Zeitlinger.

Cassady, J. C. (2001a). 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.

Cassady, J. C. (2001b). Self-Reported GPA and SAT: A Methodological Note. *Practical Assessment, Research, and Evaluation, 7(12).* Available online: http://pareonline.net/getvn.asp?v=7&n=12.

Cassady, J. C. (2004a). The influence of cognitive test anxiety across the learning-testing cycle. *Learning and Instruction, 14*(6), 569–592.

Cassady, J. C. (2004b). The impact of cognitive test anxiety on text comprehension and recall in the absence of salient evaluative pressure. *Applied Cognitive Psychology, 18*(3), 311–325.

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety, procrastination, and academic performance. *Contemporary Educational Psychology, 27,* 270–295.

Cassady, J. C., Mohammed, A., & Mathieu, L. (2004). Cross-cultural differences in test anxiety: Women in Kuwait and the United States. *Journal of Cross-Cultural Psychology, 35*(6), 715–718.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling, 13*(3), 440–464.

Everson, H. T., Smodlaka, I., & Tobias, S. (1995). Exploring the relationship of test anxiety and metacognition on reading test performance: A cognitive analysis. *Anxiety, Stress, and Coping, 7,* 85–96.

Ferrando, P., Verea, M., & Lorenzo, U. (1999). Evaluación Psicométrica del Cuestionario de Ansiedad y Rendimiento en una muestra de escolares. *Psicothema, 11*(1), 225–236.

Flett , G. L., Blankstein, K. R., & Boase, P. (1987). Self-focused attention in test anxiety and depression. *Journal of Social Behaviour and Personality, 2* (2), 259–266.

Furlan, L. (2006). Ansiedad ante los examenes: que se evalua y como? *Evaluar, 6,* 32–51.

Gutierrez Calvo, M., & Avero, P. (1995) Ansiedad, estrategias auxiliares y comprensión lectora: déficit de procesamiento vs falta de confianza. *Psicothema, 7*(3), 569–578.

Gutierrez Calvo, M. (1996). Ansiedad y Deterioro Cognitivo: Incidencia en el rendimiento académico. Rev. *Ansiedad y Estrés, 2*(2–3), 173–194.

Hembree, R. (1988). Correlates, causes, and treatment of test anxiety. *Review of Educational Research, 58,* 47–77.

Hocevar, D., & El-Zahhar, N. (1992). Cross-cultural differences in test anxiety: Establishing transliteral equivalence. In K. A. Hagtvet & T. B. Johnsen (Eds.), *Advances in test anxiety research* (Vol. 7, pp. 48–61). Lisse, The Netherlands: Swetts & Zeitlinger.

Hoddap, V. (1996). The TAI-G: A multidimensional approach to the assessment of test anxiety. In C. Schwarzer & M. Zeidner (Eds.), *Stress, Anxiety, and Coping in Academic Settings* (pp. 95–130). Tübingen: Francke.

Hoddap, K. A., & Benson, J. (1997). The motive to avoid failure and test anxiety responses: Empirical support for integration of two research traditions. *Anxiety, Stress and Coping, 10,* 35–57.

Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. Hoyle (Ed*.), Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.

Keith, N., Hoddap, V., Shermelleh–Engel, K., & Mossbrugger, H. (2003). Cross sectional and longitudinal confirmatory factor models for the Germany Test Anxiety Inventory: A construct validation. *Anxiety, Stress and Coping, 16*(3), 251–270.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: a distingtoion and some initial data. *Psychological Reports, 20,* 975–978.

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of Abnormal and Social Psychology, 47,* 166–173.

Morris, L. W., Davis, M. A., & Hutchings, C. H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale. *Journal of Educational Psychology, 73*, 541–555.

Musch, J., & Bröder, A. (1999). Überprüfung der psychometrischen Eigenschaften und der Validität des multidimensionalen Prüfungsängstlichkeitsinventars TAI-G [Investigating the psychometric properties and the validity of the multidimensional German Test Anxiety Inventory TAI-G]. *Zeitschrift für Pädagogische Psychologie, 13,* 100–105.

Naveh-Benjamin, M., McKeachie, W. J., & Lin, Y. (1987). Two types of test-anxious students: Support for an information processing model. *Journal of Educational Psychology, 79,* 131–136.

Rost, D. H., & Schermer, F. J. (1992). "eactions to Tests" (RTT) and "Manifestations of Test Anxiety" (DAI-MAN): One and the same concept? In K. Hagtvet &T. Backer Johnsen (Eds.), *Advances in Test Anxiety Research* (Vol. 7, pp. 114–129). Lisse/Berwyn: Swets&Zeitlinger/Swets North America.

Sarason, I. G. (1978). The Test Anxiety Scale: concept and research. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and Anxiety* (Vol. 5, pp 193–216). Washington, DC: Hemisphere.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology, 46,* 929–938.

Schutz, P. A., & Davis, H. A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist, 35,* 243–256.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *State-trait anxiety inventory: STAI.* Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gonzalez, H. P., Taylor, C. J., Anton, W. D., Algaze, B., Ross, G. R., & Westberry, L. G. (1980). *Test Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Valero Aguayo, L. (1999). Evaluación de ansiedad ante exámenes: datos de aplicación y fiabilidad de un cuestionario CAEX. *Anales de Psicología, 15*(2), 223–231.

Vijver, F J. R. van de, & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology, 47*(4), 263–280.

Ware, W. B., Gallasy, J. P., & Dew, K. H. (1990). The Test Anxiety Inventory: A confirmatory factor analysis. *Anxiety Research, 3,* 205–212.

Zohar, D. (1998). An additive model of test anxiety: Role of exam-specific expectations. *Journal of Educational Psychology, 90,* 330–340.