

IS THE TEST ATTITUDE SURVEY PSYCHOMETRICALLY SOUND?

JULIE M. MCCARTHY
University of Toronto

RICHARD D. GOFFIN
University of Western Ontario

Research in the realm of selection and educational testing has begun to explore test takers' attitudes toward tests. However, valid measurement of test-taking attitudes is required before concrete conclusions may be drawn. The purpose of this investigation was to examine the psychometric properties of a popular measure of test-taking attitudes that may be used in occupational or educational selection testing applications—the Test Attitude Survey (TAS). Participants included 242 undergraduate students, each of whom completed a battery that included personality and test attitude items. Although results indicated adequate convergent validity for scores on the TAS, there is room for improvement to the overall structural validity of scores on this measure.

Keywords: test anxiety; test attitudes; selection; Test Attitude Survey (TAS)

Standardized tests are frequently used to make selection decisions in occupational and educational settings (Sackett, Schmitt, Ellingson, & Kabin, 2001). When carefully constructed, these tests can be impressive in their abil-

We are grateful to Ryan Stanga and Andre Popadynech for their assistance in the collection and entry of data. A preliminary version of this article was presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, New Orleans, LA, April 2000. This work was supported by Research Grant No. 410-99-0247 from the Social Sciences Research Council of Canada to the second author. Correspondence concerning this article should be sent to Julie M. McCarthy. Correspondence concerning this article should be sent to Julie M. McCarthy, Division of Management and Joseph L. Rotman School of Management, University of Toronto, 1265 Military Trail, Scarborough, Ontario, Canada, M1C 1A4; e-mail: mccarthy@utsc.utoronto.ca.

Educational and Psychological Measurement, Vol. 63 No. 3, June 2003 446-464
DOI: 10.1177/0013164402251038
© 2003 Sage Publications

ity to predict future performance (Schmidt & Hunter, 1998). Nevertheless, it is common for test takers to report unfavorable attitudes, such as high levels of anxiety, toward selection instruments (Schmit & Ryan, 1997). Test-taking attitudes have been the focus of recent attention in the realm of personnel selection (e.g., Arvey, Strickland, Drauden, & Martin, 1990; Birenbaum & Gutvirtz, 1993; Sarason, 1984; Schmit & Ryan, 1992; Zeidner, 1991). According to this new line of research, test takers' attitudes toward tests can be perceived as individual difference variables and may include such concepts as test-taking anxiety, test-taking motivation, lack of concentration, worry, test ease, bodily symptoms, and task-irrelevant thinking (Arvey et al., 1990; Sarason, 1984; Zeidner, 1991).

The importance of examining the test-taking attitudes of job applicants and students is highlighted by three key points. First, test-taking attitudes may be related to actual test performance. It is believed that extreme levels of test-taking attitudes (e.g., high levels of anxiety and/or low levels of motivation) may have detrimental effects on test performance, resulting in scores that may not accurately reflect an individual's true abilities (Arvey et al., 1990). In support of this hypothesis, several investigations have found a positive relation between test-taking motivation and test performance (Arvey et al., 1990; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Sanchez, Truxillo, & Bauer, 2000). Evidence also suggests that test-taking anxiety is negatively related to test performance (Arvey et al., 1990; Schmit & Ryan, 1992). Further investigation of test-taking attitudes is advantageous, as it may ultimately lead to the development of approaches for the effective treatment of individuals with debilitating test-taking attitudes.

Second, test-taking attitudes may influence the validity of selection tests. This proposition is linked to the first point and asserts that if attitudes toward tests are related to test performance, then to the extent that they result in less accurate measurement of a person's level on the construct of interest, they may ultimately result in the selection of a less promising set of individuals. This possibility is supported by the results of an investigation conducted by Schmit and Ryan (1992), who found that test-taking attitudes moderated the validity of cognitive ability and personality-based selection instruments. The organizational implication is clear—if test-taking attitudes systematically mask or exaggerate an applicant's true likelihood of success in terms of the criterion that is being predicted (e.g., job performance or scholastic success), then measures of test-taking attitudes should be incorporated into selection procedures.

Third, variations in test-taking attitudes across different groups may contribute to racial differences in test performance (Arvey et al., 1990; Chan et al., 1997; Chan, Schmitt, Sacco, & DeShon, 1998). This is a critical issue, as racial differences in scores on standardized tests are well documented and often lead to adverse impact in occupational and educational selection contexts (Sackett et al., 2001). If racial differences in test-taking attitudes are par-

tially responsible for the occurrence of adverse impact, it is conceivable that these differences might be amenable to statistical control, provided that test-taking attitudes can be appropriately measured. The development of treatments for improving the test-taking attitudes of certain groups may also be advantageous.

For the reasons just discussed, test-taking attitudes appear to be an important set of constructs to assess in selection-testing contexts. It is therefore surprising that relatively few investigations of test-taking attitudes have been conducted (Ryan & Ployhart, 2000). This may be due, in part, to the lack of well-established measurement tools. Indeed, before empirical research can provide concrete information about the role of applicant and student attitudes toward tests, measures of test-taking attitudes that can yield valid scores are required.

The most common measure of test-taking attitudes used in selection testing applications is the Test Attitude Survey (TAS), which was developed by Arvey et al. in 1990. The TAS is a multidimensional instrument composed of the following nine subscales: Motivation (10 items), Lack of Concentration (4 items), Belief in Tests (4 items), Comparative Anxiety (10 items), Test Ease (4 items), External Attribution (5 items), General Need Achievement (3 items), Future Effects (3 items), and Preparation (2 items). These scales cover a broad spectrum of test attitudes and have the potential to provide detailed insight into a test taker's frame of mind. To date, at least nine published investigations have used the TAS as a measure of test-taking attitudes (see Arvey et al., 1990; Chan et al., 1997, 1998; Neuman & Baydoun, 1998; Ryan, Ployhart, Greguras, & Schmit, 1998; Sanchez et al., 2000; Schmit & Ryan, 1992, 1997; Schmitt, Chan, Sacco, McFarland, & Jennings, 1999). In addition, the Social Science Citation Index indicates that more than 40 published studies have cited the Arvey et al. article.

Unfortunately, little research has examined the psychometric properties of the TAS. In fact, only the authors of the TAS (Arvey et al., 1990) have systematically examined both the reliability and validity of the scores on their instrument. In their 1990 article, they describe initial reliability and validity estimates for TAS scores, as well as the procedures used to develop the TAS scales. Internal consistency reliabilities (Cronbach's alpha) were computed for the scores on each of the nine scales, and were found to range from .56 to .85. Similar results were reported by Schmit and Ryan (1992), who found subscale score alpha coefficients ranging from .50 to .94. It is noteworthy that each of these investigations reported score reliability estimates lower than .60 for the scores on the following scales: Test Ease, External Attribution, General Need Achievement, and Future Effects. Murphy and Davidshofer (1991) strongly recommend that internal consistency reliability for scores should exceed .60.

Arvey et al. (1990) also examined the construct validity of scores on five of the TAS scales (Test Ease, Comparative Anxiety, External Attributions, Motivation, and Lack of Concentration). The results confirmed the prediction that subjects who were given the easier test demonstrated higher scores on the Test Ease scale. Also consistent with predictions, individuals who were administered the difficult test demonstrated higher scores on the Comparative Anxiety and External Attribution scales.

A separate investigation was performed to examine the construct validity of scores on the Motivation and Lack of Concentration scales (Arvey et al., 1990). Consistent with predictions, the computerized version was related to higher scores on the Motivation and Concentration scales. Although these results provide some evidence of the construct validity of subscale scores, some unexpected results were also found. Specifically, the authors found it difficult to explain why participants in the paper-and-pencil condition reported higher scores on the Belief in Tests, External Attributions, Need for Achievement, and Future Effects scales. Based on the aforementioned results, Arvey et al. (1990) concluded that the validity of scores on the TAS was "sufficiently adequate" (p. 703) to justify its use in empirical research.

The Present Study

As described, there is clear need for valid measures of test-taking attitudes that can be applied to educational and pre-hiring selection testing contexts. The TAS has the potential to fill this niche, as it is a multidimensional measure of test-taking attitudes that is suitable for use in selection-testing applications. To this end, research is required to assess whether the scales comprising the TAS need to be further refined or replaced. Although the aforementioned studies are important first steps, a more comprehensive analysis of the psychometric properties and construct validity of scores on the TAS is essential, particularly in light of the fact that the TAS has been used in at least nine empirical investigations. Accordingly, the purpose of this study was to investigate the psychometric properties of the TAS by conducting a step-by-step examination using both univariate and multivariate statistical analyses.

Method

Participants and Procedure

A total of 248 undergraduate students from the University of Western Ontario participated in the present investigation (145 women, 103 men; aged 17 to 49 years). Each individual completed a battery of test attitude, personality, and cognitive ability items as part of a larger program of research.

Participants were tested in groups ranging from 5 to 20 individuals. On arrival, each individual was provided with detailed instructions for completing the paper-and-pencil test battery. The test battery was composed of three sections: Section 1, attitude and personality measures; Section 2, a simulated selection test consisting of cognitive ability and personality measures; and Section 3, items from the TAS (Arvey et al., 1990). Verbal instructions were provided at the beginning of each section.

After participants had completed Section 1, the examiner proceeded to Section 2, in which participants were asked to play the role of an applicant to a prestigious university who must obtain an impressive score on a selection test to gain admission (e.g., Schmit & Ryan, 1992). This scenario was deemed appropriate because the majority of students are aware of the qualifications required to be a successful university student. To make this part of the session seem more realistic, cash awards of \$20.00 were provided for the top scores on the tests in lieu of a job, and participants were informed of this contingency. The cash awards were used to make the simulation more like a true selection scenario where rewards (i.e., getting hired or getting accepted into an elite university) are, indeed, associated with superior test performance. Previous use of a similar procedure by Schmit and Ryan (1992) had shown promise. Prior to the administration of the TAS in Section 3, participants were informed that the selection simulation was over and that no prizes would be associated with their responses to the TAS items. On completion of Section 3, the examiner collected the test booklets and gave each participant a debriefing form and a research credit sheet.

Measures

Test attitudes. Test attitudes were measured with the TAS (described earlier; see Arvey et al., 1990). The response format was a 5-point Likert scale ranging from 1 (*strongly agree*) to 5 (*strongly disagree*). The scales and individual items contained in this measure are presented in Table 3.

Personality. Four experts in the fields of personality and industrial/organizational psychology were recruited and asked to provide judgments as to which personality traits bore the most resemblance to test-taking attitudes as defined in the TAS. The selected traits were later used to evaluate aspects of construct validity. The traits were selected from a prechosen list of Jackson Personality Inventory–Revised (JPI-R) (Jackson 1994), Personality Research Form (PRF-E) (Jackson, 1987), and Basic Personality Inventory (BPI) (Jackson, 1986) scales. Experts identified the 10 dimensions that they felt were most strongly related to test-taking attitudes. The dimensions were listed in rank-order format, with the most relevant scale receiving a score of

10 and the least relevant scale receiving a score of 1. Average relevance scores across the four judges were computed for each dimension.

Due to temporal constraints, the maximum number of dimensions included in the questionnaire was limited to 6. Selection of the final dimensions was based on a two-step process. First, dimensions were eliminated if they were not included in at least 2 of the judges' top 10 rankings. Second, the 6 dimensions with the highest average relevance scores were identified. These 6 dimensions included the Anxiety (20 items; $\alpha = .82$; i.e., nervousness and tension) and Organization (20 items; $\alpha = .79$; i.e., orderliness and tidiness) subscales of the JPI-R; the Self-Depreciation (18 items; $\alpha = .61$; i.e., self-critical thoughts) and Hypochondriasis (20 items; $\alpha = .80$; i.e., preoccupation with health-related concerns) subscales from the BPI; and the Achievement (20 items; $\alpha = .57$; i.e., motivation to succeed) and Endurance (16 items; $\alpha = .75$; i.e., stamina and perseverance) subscales from the PRF-E. The PRF-E Social Desirability scale (19 items; $\alpha = .68$) was also included due to the prominence of this construct in the assessment of individual differences and its importance in the assessment of discriminant validity (Helmes, 2000). Consistent with the TAS, the response format for each of the aforementioned personality dimensions was a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Infrequency. To detect nonpurposeful or random responding, the Infrequency scale of the PRF-E (Jackson, 1987) was administered to participants. This scale consists of 16 items and uses the 5-point Likert scale described earlier. Jackson (1987) reported an internal consistency reliability of .71 for the scores on this scale using a college sample.

Analytic Strategy

A predetermined series of statistical procedures were followed in the assessment of the TAS. In the first step, the most basic psychometric properties of the TAS items and scale scores were examined (e.g., descriptive statistics and internal consistency reliability estimates).

Next, an assessment of the convergent and discriminant validity of overall scores on the TAS was conducted by computing correlations between scores on the overall TAS and relevant personality traits that were selected via expert judgment. Examining the construct validity of overall scores on the TAS was deemed appropriate, as previous researchers have used such a score (i.e., Neuman & Baydoun, 1998; Schmit & Ryan, 1992) to capture a general positive test-taking attitude. In addition to the overall TAS, the convergent and discriminant validity of the Motivation and Comparative Anxiety scale scores were also examined due to the recent attention surrounding these con-

structs (Arvey et al., 1990; Birenbaum & Gutvirtz, 1993; Chan et al., 1997; Sarason, 1984; Zeidner, 1991).

In the next step, the factor structure of scores on the TAS was examined using confirmatory factor analyses. Factor structure is a crucial aspect of construct validity, and refers to the degree to which the empirical factor structure of the scores on a measure are congruent with the presumed theoretical structure of the scores (Loevinger, 1967). The goal of these analyses was to test the multidimensional TAS structure presumed by Arvey et al. (1990) and the unidimensional TAS structure presumed by Schmit and Ryan (1992). To avoid biases resulting from the Future Effects and Preparation scales (which were not directly applicable to the research paradigm employed), these scales were excluded from the analyses. Accordingly, an oblique seven-factor model was tested in which the factors corresponded to the TAS scales. An a priori decision was made to conduct an exploratory principal component analyses if neither the seven-factor nor the one-factor a priori models were confirmed. In the event that neither the one-factor nor the seven-factor model described the data, the exploratory analysis would at least provide a preliminary indication of the true dimensionality of the TAS. Nevertheless, the exploratory analysis would be interpreted with caution, as cross-validation with a new sample would be required before any firm conclusions could be drawn.

Results

Preliminary Analyses

Six participants were eliminated due to high infrequency scores (i.e., average item scores higher than 2.00), indicating potentially nonpurposeful responding. This resulted in a final sample of 242 participants (142 women, 100 men).

Properties of the TAS Items

Psychometrically, items with means that are either too high or too low are problematic because they result in reduced item variance due to ceiling or floor effects. Examination of item means revealed that they were acceptable, ranging from 1.44 to 4.17 on a 5-point scale. Similarly, item standard deviations indicated a reasonable level of item variability, ranging from 0.56 to 1.17.

Table 1
Descriptive Statistics, Score Alpha Coefficients, and Item-Total Correlations

TAS Scale	<i>M</i>	<i>SD</i>	α	Range of Item-Total Correlations
Motivation	3.57	0.65	.90	.55 to .76
Lack of Concentration	2.47	0.77	.76	.48 to .62
Belief in Tests	2.93	0.65	.58	.27 to .47
Comparative Anxiety	2.59	0.61	.82	.33 to .69
Test Ease	2.80	0.57	.56	.16 to .49
External Attribution	2.33	0.59	.49	.17 to .32
General Need Achievement	3.76	0.60	.49	.29 to .35
Preparation	1.49	0.73	.83	.72 to .72
Future Effects	1.72	0.74	.44	.12 to .41
Overall TAS	3.49	0.39	.87	-.27 to .69

Note. TAS = Test Attitude Survey (Arvey, Strickland, Drauden, & Martin, 1990); α = Cronbach's alpha; Overall TAS = Motivation + Lack of Concentration (reversed) + Belief in Tests (reversed) + Comparative Anxiety (reversed) + Test Ease (reversed) + External Attribution (reversed) + General Need for Achievement—the Future Effects and Preparation scales were not included.

Psychometric Properties of the TAS Scales

Means and standard deviations for each of the TAS subscale scores, as well as the overall TAS, are presented in Table 1. Estimates of internal consistency reliability were also calculated for scores on each scale. As previously noted, Murphy and Davidshofer (1991) have suggested that score alpha coefficients should exceed .60 to be considered acceptable. Unfortunately, five of the nine scale scores had alpha coefficients less than .60 (Belief in Tests, Test Ease, External Attribution, General Need Achievement, and Future Effects). These results are generally consistent with those reported by Arvey et al. (1990) and Schmit and Ryan (1992), who found score alpha coefficients lower than .60 for the latter four scales.

Item-total correlations (corrected for item overlap) are another important consideration when examining the reliability of scores and reflect the relationship between the item score and the entire subscale score. Ideally, these values should be in the positive direction and should be moderate to high in magnitude (Murphy & Davidshofer, 1991). In the present sample, the corrected item-total correlations for item scores from four of the nine scales (Motivation, Lack of Concentration, Comparative Anxiety, and Preparation) were moderate to high (see Table 1). Unfortunately, corrected item-total correlations were small to medium in strength for the item scores of the remaining five scales. As expected, these were the same five scale scores that demonstrated low alpha coefficients.

Table 2
Convergent and Discriminant Validity of Test Attitude Survey Scores

Personality Trait	TAS Overall		TAS Motivation		TAS Anxiety	
	Expected Relation	Actual Relation	Expected Relation	Actual Relation	Expected Relation	Actual Relation
BPI Anxiety	-	-.24**	-	.04	+	.49**
JPI Organization	+	.20**	+	.09	0	-.11
BPI Hypochondriasis	-	-.21**	0	.09	+	.39**
BPI Self-Depreciation	-	-.39**	-	-.20**	+	.34**
PRF Achievement	+	.24**	+	.08	-	-.15*
PRF Endurance	+	.27**	+	.05	-	-.18**
PRF Desirability	+	.27**	+	-.04	-	-.36**

Note. Overall TAS = Motivation + Lack of Concentration (reversed) + Belief in Tests (reversed) + Comparative Anxiety (reversed) + Test Ease (reversed) + External Attribution (reversed) + General Need for Achievement—the Future Effects and Preparation scales were not included; BPI = Basic Personality Inventory (Jackson, 1994); JPI = Jackson Personality Inventory (Jackson, 1976); PRF = Personality Research Form (Jackson, 1987).

* $p < .05$. ** $p < .01$.

Convergent and Discriminant Validity

The assessment of convergent and discriminant validity involved computing the correlations of the personality dimension scores (selected to be relevant to test-taking attitudes by the aforementioned panel of experts) with the TAS overall score, the TAS Motivation scale score, and the TAS Comparative Anxiety scale score. The results of these are presented in Table 2. The second column in Table 2 provides the relations predicted by the present authors between the overall TAS and each personality trait. The overall TAS was expected to be related to, but far from congruent with, each of these constructs. Accordingly, each of the predicted relations was anticipated to be of medium strength (low: .10 to .29; medium: .30 to .49; high: > .50).

As indicated in Table 2, the relation between scores on the personality construct of Anxiety and the overall TAS score was expected to be negative. The rationale for this prediction was that the TAS is an overall measure of positive test-taking attitudes that includes negatively keyed Comparative Anxiety items, positively keyed Motivation items, and positively keyed Need for Achievement items. Scores on the second personality scale, Organization, were expected to demonstrate a positive relation with the overall TAS score. This was based on the fact that individuals who are well organized in their daily activities will also be more likely to demonstrate positive test-taking attitudes. The third scale, Hypochondriasis, measures the extent to which individuals are preoccupied with physical illness and bodily concerns (Jackson, 1986). Individuals with high levels of hypochondriasis are more likely to report somatic symptoms, such as headaches and stomach pains. It was there-

fore predicted that individuals who chronically suffer from hypochondriacal tendencies would experience less positive test-taking attitudes, resulting in a negative correlation between Hypochondriasis scores and TAS scores.

The relation between scores on the fourth personality scale, Self-Depreciation, and the overall TAS scores was also expected to be negative because self-depreciation is associated with a cynical frame-of-mind, and should therefore demonstrate a negative relation with positive test-taking attitudes (as measured by the overall TAS score). Scores on the next personality scale, Achievement, were expected to demonstrate a positive relation with the overall TAS scores. It was expected that someone who is achievement-oriented would be more likely to demonstrate positive test-taking attitudes such as Motivation, Need for Achievement, and Concentration. Similarly, a positive relation was anticipated between the Endurance scale scores and the overall TAS score. This prediction was based on the fact that individuals with high endurance scores are persistent and determined when working on tasks and are therefore more likely to exhibit positive test-taking attitudes.

Scores on the final scale, Social Desirability, were ultimately expected to demonstrate a positive relation with the overall TAS score. Initially, we felt that test-taking attitudes should be relatively distinct from socially desirable responding. This belief was consistent with the Jacksonian tradition (e.g., Jackson, 1970), which holds that constructs should be relatively uncontaminated by social desirability variance. However, on further consideration, it became evident that social desirability may be an important component of test-taking attitudes (Schmit & Ryan, 1992). In support of this proposition, research has indicated that test-taking attitudes scores and social desirability scores are positively related (Schmit & Ryan, 1992). Specifically, Schmit and Ryan (1992) found that participants with high TAS scores were also identified as those with a tendency to fake good. To this end, our initial position was revised and a positive relation of medium strength was anticipated between scores on the Social Desirability scale and the overall TAS score.

The third column of Table 2 contains the observed relationships between the personality trait scores and the overall TAS scores. As illustrated, the overall TAS scores demonstrated statistically significant relationships in the hypothesized direction with each of the personality dimensions. However, the anticipated strength of relations was not confirmed. Specifically, according to the standards set by Cohen and Cohen (1975), each of these relations was found to be small (not medium) in magnitude. One exception was the relation between the overall TAS score and the Self-Depreciation scale score, which was found to be medium in strength.

Correlations were also obtained between the scores on the personality scales and the Motivation scale of the TAS. As indicated in the fourth column of Table 2, it was anticipated that the Motivation scale scores would demonstrate positive relations of medium strength with the Organization, Achieve-

ment, and Endurance scale scores, as individuals who are motivated to succeed are more likely to be organized, goal-oriented, and persistent. A positive relation of medium strength between Motivation and Desirability scale scores was also anticipated, as highly motivated individuals are more likely to fake good on assessment instruments (Schmit & Ryan, 1992). Based on the findings of Arvey et al. (1990), a negative relation of medium strength was anticipated between the TAS Motivation scale scores and the Anxiety scale scores. It was also expected that scores on the Motivation scale would demonstrate a negative relation of medium strength with the Self-Depreciation scale scores, as an individual motivated to do well is less likely to engage in self-depreciatory behaviors. No relation was anticipated between the Motivation scale scores of the TAS and the Hypochondriasis scale scores, as these constructs appear to be conceptually distinct.

Examination of Table 2 reveals that only one of the a priori predictions involving the TAS Motivation scale scores was confirmed. Specifically, as anticipated, no relation was found between scores on the Motivation and Hypochondriasis scales ($r = .09, ns$). A negative relation was found between scores on the TAS Motivation scale and scores on the Self-Depreciation scale; however, contrary to expectations, this relation was small rather than medium in magnitude ($r = -.20, p < .01$). The five remaining relations involving the TAS Motivation scores were extremely small in magnitude, ranging from $-.04$ to $.09$, and were inconsistent with the a priori hypotheses.

Relations between scores on the personality scales and scores on the TAS Anxiety scale were also computed. Not surprisingly, the relation between the TAS Anxiety scale scores and the BPI Anxiety scale scores was expected to be positive and medium to large in strength. Positive relations of medium strength were anticipated between TAS Anxiety scores and two of the personality scales: Hypochondriasis and Self-Depreciation. Scores on Hypochondriasis were expected to be related to TAS Anxiety scores because individuals with high levels of anxiety are more likely to report hypochondriac-related symptoms (i.e., fast heartbeat, loss of breath, headaches) than individuals with low levels of anxiety (Sarason, 1984). The positive relation between TAS Anxiety scores and Self-Depreciation scores was anticipated based on the fact that anxiety is often associated with a fear of failure (Liebert & Morris, 1967; Spielberger, 1980), and individuals who anticipate personal failure are more likely to experience negative self-thoughts (i.e., self-depreciation).

Negative relations of medium strength were anticipated between TAS Anxiety scores and the Achievement and Endurance scores because the items on the latter two scales generally reflect a positive attitude regarding one's capabilities, whereas the items on the former scale are reflective of self-doubts. A negative relation of medium strength was also expected between scores on the TAS Anxiety scale and scores on the Social Desirability scale

because the items of the former scale involve admissions of self-doubts and weaknesses, whereas the Social Desirability scale reflects a tendency to respond in a more positive, self-affirming way that others would generally approve of. Finally, no relation was expected between scores on TAS Anxiety and scores on the Organization scale, as these constructs appear to be conceptually distinct.

As indicated in Table 2, five of the a priori predictions for TAS Anxiety scores were confirmed. To be specific, the relation between scores on TAS Anxiety and BPI Anxiety was positive and large ($r = .49, p < .01$), the relation between TAS Anxiety scores and Hypochondriasis scores was positive and of medium strength ($r = .39, p < .01$), the relation between TAS Anxiety scores and Self-Depreciation scores was positive and of medium strength ($r = .34, p < .01$), the relation between scores on TAS Anxiety and Desirability was negative and medium in strength ($r = -.36, p < .01$), and the relation between TAS Anxiety scores and Organization scores was nonsignificant ($r = .11, ns$). Finally, the relations between scores on TAS Anxiety and scores on the Achievement ($r = -.15, p < .05$) and Endurance ($r = -.18, p < .05$) scales were found to be in the anticipated direction but were small, whereas medium strength relations were expected.

Overall, evidence for the convergent validity of the scores on the overall TAS and the Anxiety scale of the TAS was modest. Although the majority of relations were found to be in the predicted direction, the strength of several observed relations was not consistent with a priori predictions. Even less support was obtained for the convergent validity of the TAS Motivation scale. In terms of discriminant validity, the relation between the TAS Anxiety scale scores and the Desirability scale scores was moderate. This finding was anticipated a priori and is not unlike results obtained with many individual difference measures (see Helmes, 2000). Inconsistent with predictions, small relations were found between the overall TAS score and the Desirability score ($r = .27, p < .01$), as well as between the TAS Motivation score and the Desirability score ($r = -.04, ns$). However, the magnitude of the relation between the overall TAS score and the Desirability scale score approached medium strength (as expected). In sum, we judged the discriminant validity of only the TAS overall score and the TAS Anxiety score to be adequate.

Confirmatory Factor Analyses

The confirmatory factor analyses were conducted on the variance-covariance matrices using LISREL 8.20 (Joreskog & Sorbom, 1997). Maximum likelihood estimation was used, and three indices were used to assess the fit of the models: the chi-square index, the Root Mean Square Error of Approximation (RMSEA), and the Relative Noncentrality Index (RNI). Neither the one-factor nor the seven-factor model provided an acceptable fit to

the data—one-factor model: $\chi^2(780) = 5,638$, RMSEA = .25, RNI = .00; seven-factor model: $\chi^2(740) = 2,582$, RMSEA = .11, RNI = .62. Similarly, poor fits were obtained when the factors were constrained to be orthogonal.

Exploratory Principal Component Analyses

The discouraging results obtained in the confirmatory factor analyses necessitated the use of exploratory techniques to further investigate the underlying structure of the TAS. Principal component extraction with a varimax rotation was used. To be consistent with the confirmatory analyses, the Future Effects and Preparation scales were excluded from these analyses.

A combination of eigenvalue and Scree plot data was used to determine the appropriate number of components to extract. The Scree plot indicated that a three component model would be the most appropriate fit to the data. All three of these components had eigenvalues greater than one, but they collectively accounted for only 41% of the variance. Examination of the Scree plot also suggested that a seven component solution was possible; seven components accounted for 57% of the variance and all had eigenvalues greater than one. Accordingly, two principal component analyses were conducted; the first extracted three components, and the second extracted seven.

Three component model. The results of the three component model are presented in Table 3. The items with high pattern/structure coefficients on the first component generally measured test-taking motivation. An exception appears for the item, “I found this test or tests interesting and challenging,” which was presumed by Arvey et al. (1990) to be measuring Test Ease but has an apparently irrelevant high pattern/structure coefficient on Component 1. The second component contains items measuring test-taking anxiety and test difficulty, both of which are clearly reflective of self-doubts with regard to testing. Therefore, the most appropriate label for this component is “Self-Doubts Regarding Test-Taking.” An examination of the third component reveals that it is composed of items tapping dislike of tests, inability to do well on tests, and lack of concentration in test-taking situations. Both an inability to do well on tests and a lack of concentration on tests may result from an overall aversion to the testing process. Accordingly, an appropriate label for this component is “Dislike of Tests.”

Seven component model. Results of this analysis indicated that all of the Motivation items had high pattern/structure coefficients on Component 1, all of the Lack of Concentration items had high pattern/structure coefficients on Component 3, all of the Belief in Tests items had high pattern/structure coefficients on Component 6, and all of the General Need for Achievement items had high pattern/structure coefficients on Component 7. However, items

from the remaining three scales (Comparative Anxiety, External Attribution, and Test Ease) were scattered across the seven components. Specifically, the Comparative Anxiety items were dispersed across Components 2, 3, 4, and 5; the External Attribution items were dispersed across Components 2, 3, 5, and 6; and the Test Ease items were dispersed across Components 1, 2, and 5. It was not possible to interpret and label these components. Results of the seven component model are not presented.

Discussion

Findings revealed that scores had adequate levels of alpha reliability and reasonably high item-total correlations for four of the TAS subscales (i.e., Motivation, Lack of Concentration, Comparative Anxiety, and Preparation), as well as the overall TAS. The TAS subscale scores demonstrating unacceptably low levels of alpha reliability and low item-total correlations were the Belief in Tests, Test Ease, External Attribution, General Need Achievement, and Future Effects scales. Results also indicated modest support for the convergent and discriminant validity of scores on the overall TAS, as well as the TAS Anxiety scale. In contrast, evidence for the convergent and discriminant validity of the scores on the TAS Motivation scale was weak. Finally, findings suggested that there is room for improvement to the overall structural validity of TAS scores. Several elements were considered in reaching this conclusion. First, the confirmatory factor analysis failed to replicate the theoretical structure assumed to underlie the TAS. Second, the items comprising the TAS scales were dispersed among the components in the exploratory factor analysis in a manner that was not consistent with the scoring key. Third, the exploratory factor analysis suggested a three-component solution, which differs substantially from the seven dimensions and scoring key underlying the respective items.

At first glance, these results appear counterintuitive—the psychometric properties of the individual scales are satisfactory, but the structural validity of scores on the TAS is weak. What this implies, however, is that there may be a more advantageous way to group the items that would yield an improved measure of test-taking attitudes. The findings of the current study suggest that the most supportable structure may be a three-factor model (Motivation, Dislike of Tests, Self-Doubts Regarding Test-Taking). Cross-validation is required to properly assess this structure. If subsequent investigations confirm the three-factor model, then regrouping the items based on the new structure will enhance the psychometric properties and structural validity of the TAS. However, on conceptual grounds, the veracity of the three-factor model is questionable because it combines Test-Taking Anxiety and Test Difficulty into a single factor. Given that Test Difficulty is strongly influenced by factors that are external to the test taker (Arvey et al., 1990), whereas Test-

Table 3
Principal Components Analysis—Rotated Three Factor Solution

Scale	Item	Component 1	Component 2	Component 3
Motivation	I tried to do the very best I could on this test.	0.78	-0.03	-0.11
	I wanted to do well on this test.	0.77	-0.11	0.08
	I tried my best on this test.	0.74	-0.07	-0.12
	I didn't put much effort into this test. ^a	0.73	-0.04	-0.18
	While taking this test, I concentrated and tried to do well.	0.72	-0.05	-0.35
	I was extremely motivated to do well on this test.	0.70	0.15	-0.15
	I pushed myself to work hard on this test.	0.69	0.11	-0.13
	I just didn't care how I did on this test. ^a	0.69	0.13	-0.31
	Doing well on this test is important to me.	0.67	0.13	0.00
	I want to be among the top scorers on this test.	0.63	-0.02	0.01
Test Ease	I found this test interesting and challenging. ^a	-0.58	-0.26	0.25
General Need Achievement	I try to do well in everything I undertake.	0.31	-0.09	-0.12
	Once I undertake a task, I usually push myself to my limits.	0.19	-0.04	-0.18
Belief in Tests	This test was a good reflection of what a person could do in the job. ^a	-0.26	0.05	-0.02
Comparative Anxiety	During the testing, I often thought about how poorly I was doing.	-0.09	0.70	0.16
	During the testing, I got so nervous I couldn't do as well as I should have.	0.09	0.66	0.34
	I probably didn't do as well as most of the other people who took the test.	-0.09	0.58	0.35
	During the test, I found myself thinking of the consequences of failing.	0.16	0.50	0.36
	I usually get very anxious about taking tests.	0.09	0.48	0.29
	I expect to be among the people who score really well on this test. ^a	-0.45	0.47	0.00

Test Ease	This test was too easy for me.	0.13	-0.69	0.28
	I felt frustrated because many of the questions were too difficult. ^a	-0.03	-0.69	-0.14
External Attribution	I found this test too simple.	0.07	-0.62	0.40
	I felt a lot of pressure when taking this test.	0.14	0.65	0.16
Lack of Concentration	While taking this test, I was preoccupied with how much time I had left.	0.08	0.58	0.20
	I found myself losing interest and not paying attention to the test.	-0.55	0.02	0.54
	It was hard for me to keep my mind on this test.	0.25	0.30	0.51
Belief in Tests	During the test session, I was bored.	0.42	-0.23	0.46
	I get distracted when taking tests of this type.	-0.44	0.22	0.41
	I don't believe that tests are valid.	-0.04	0.01	0.54
	This kind of test should be eliminated.	-0.32	-0.02	0.46
Comparative Anxiety	Tests are a good way of selecting people into jobs. ^a	-0.10	-0.01	0.13
	I am not good at taking tests.	0.10	0.48	0.61
	My test scores don't usually reflect my true abilities.	0.01	0.21	0.54
External Attribution	I usually do pretty well on tests. ^a	0.16	0.26	0.50
	I very much dislike taking tests of this type.	-0.32	0.15	0.46
	I became fatigued and tired during the testing.	-0.24	0.19	0.57
	I have not been feeling well lately, and this affected my performance on the test.	-0.11	0.21	0.38
General Need Achievement	The questions on this test were ambiguous and unclear.	-0.14	0.00	0.35
	In general, I like to work just hard enough to get by. ^a	0.07	-0.11	-0.27

Source. Reproduced with permission from the publisher, *Personnel Psychology*, and the authors, Arvey, Strickland, Drauden, and Martin (1990).

Note. Boldface type indicates which component each item belongs to. Scales and items are from the Test Attitude Survey (TAS). Component 1 = Motivation, Component 2 = Self-Doubts Regarding Test-Taking, and Component 3 = Dislike of Tests.

a. Indicates a negatively keyed item.

Taking Anxiety already has a venerable history as a more internal, trait-like concept (e.g., Liebert & Morris, 1967; Sarason, 1984; Spielberger, 1980), the melding of these two concepts requires further empirical support.

It is important to note that in collecting our data, considerable effort was made to ensure that the selection simulation was as realistic as possible and that the sample was comparable to actual applicants. First, an attempt was made to engender a level of test-taking motivation in our participants. Second, the context was one that the student sample was very familiar with and experienced at—taking a selection test to gain admission to a university. Moreover, nonpurposeful responding was examined by including an infrequency scale in the questionnaire. Participants who exhibited nonpurposeful responding were removed from the sample. For these reasons, the findings of the current investigation are particularly relevant to selection testing in educational settings. Caution must be applied, however, before the current findings are generalized to broader organizational contexts.

From the perspective of potential TAS users, the results of the current investigation have several important implications. First, the Anxiety scores consistently emerged as the strongest dimension of the TAS (e.g., strong item-total correlations, a high alpha coefficient, and adequate convergent and discriminant validity). However, as discussed, Anxiety scale items demonstrated high pattern/structure coefficients on a combined Test Anxiety/Test Difficulty component which we labeled “Self-Doubts Regarding Test-Taking.” The Motivation scale scores also demonstrated high item-total correlations and a high alpha coefficient. However, limited evidence for the convergent and discriminant validity of the Motivation scale scores was obtained. This is surprising, as deductive scale development procedures were used to create this scale (Arvey et al., 1990), and these procedures have been associated with several potential advantages (see Goffin, Rothstein, & Johnston, 2000). The Preparation scale scores also demonstrated strong item-total correlations and a high alpha coefficient. The psychometric properties of the six remaining scales (Belief in Tests, Test Ease, External Attribution, General Need Achievement, Future Effects, and Lack of Concentration) were not as strong.

These findings imply that researchers interested in using the TAS should focus their attention on the Anxiety scale. For the Motivation and Preparation scales, evidence of the convergent/discriminant validity of these scales is necessary before practical recommendations are made. Results also indicate that caution is warranted for the subscales of the Belief in Tests, Test Ease, External Attribution, General Need Achievement, Future Effects, and Lack of Concentration. Moreover, it is inadvisable to use an overall TAS score, as the single factor TAS solution did not provide an acceptable fit to the data.

As previously mentioned, one potential use of measures of test-taking attitudes might be to correct scores to increase the criterion-related validity of

selection procedures. In some ways this is similar to the correction for faking in personality testing (Christiansen, Goffin, Johnston, & Rothstein, 1994). To realize this goal, however, measures of test-taking attitudes that demonstrate strong psychometric properties are required. Unfortunately, test-taking attitude instruments that meet these standards are not currently available. Accordingly, the use of test-taking attitude measures should be limited to research purposes only.

In conclusion, the measurement of test-taking attitudes is an important endeavor that is gaining increased attention in organizational contexts. This is not surprising, as test-taking attitudes have the potential to play an important role in affecting the test-taking results (see Schmit & Ryan, 1992). Consequently, further consideration of test-taking attitudes could lead to enhanced utility of current selection methods and fairer treatment of individuals for whom selection tests are currently underpredicting job performance. To realize such benefits, however, reliable and valid measurement is required. Arvey et al. (1990) have provided a valuable starting point in this regard. Nevertheless, additional instrument refinement and/or development that is based on clearly defined test-attitude constructs is required to further improve the measurement of test-taking attitudes.

References

- Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test-taking. *Personnel Psychology, 43*, 695-716.
- Birenbaum, M., & Gutvirtz, Y. (1993). The relationship between test anxiety and the seriousness of errors in algebra. *Journal of Psychoeducational Assessment, 11*, 12-19.
- Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology, 82*, 300-310.
- Chan, D., Schmitt, N., Sacco, J. M., & DeShon, R. P. (1998). Understanding pretest and posttest reactions to cognitive ability and personality tests. *Journal of Applied Psychology, 83*, 471-485.
- Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47*, 847-860.
- Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Goffin, R. D., Rothstein, M. G., & Johnston, N. G. (2000). Personality and job performance: Are personality tests created equal? In R. D. Goffin and E. Helmes (Eds.) *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 249-264). Norwell, MA: Kluwer Academic Publishers.
- Helmes, E. (2000). The role of social desirability in the assessment of personality constructs. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy*. Norwell, MA: Kluwer Academic.
- Jackson, D. N. (1970). A sequential system for personality scale development. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 2, pp. 61-96). New York: Academic Press.
- Jackson, D. N. (1986). *Basic personality inventory*. Port Huron, MI: Sigma Assessment Systems.

- Jackson, D. N. (1987). *Personality Research Form—E*. Port Huron, MI: Sigma Assessment Systems.
- Jackson, D. N. (1994). *Jackson Personality Inventory—Revised*. Port Huron, MI: Sigma Assessment Systems.
- Joreskog, K. G., & Sorbom, D. (1997). *LISREL 8.20: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software.
- Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*, 975-978.
- Loevinger, J. (1967). Objective tests as instruments of psychological theory. In D. N. Jackson & S. Messick (Eds.), *Problems in human assessment* (pp. 78-123). New York: McGraw-Hill.
- Murphy, K. R., & Davidshofer, C. O. (1991). *Psychological testing: Principles and applications*. Englewood Cliffs, NJ: Prentice Hall.
- Neuman, G. A., & Baydoun, R. (1998). An empirical examination of overt and covert integrity tests. *Journal of Business and Psychology, 13*, 65-79.
- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management, 26*, 565-606.
- Ryan, A. M., Ployhart, R. E., Greguras, G. J., & Schmit, M. J. (1998). Test preparation programs in selection contexts: Self-selection and program effectiveness. *Personnel Psychology, 51*, 65-79.
- Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist, 56*, 302-318.
- Sanchez, R. J., Truxillo, D. M., & Bauer, T. N. (2000). Development and examination of an expectancy-based measure of test-taking motivation. *Journal of Applied Psychology, 85*, 739-750.
- Sarason, I. G. (1984). Stress, anxiety, and cognitive inference: Reactions to tests. *Journal of Personality and Social Psychology, 46*, 929-938.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.
- Schmit, M. J., & Ryan, A. M. (1992). Test-taking dispositions: A missing link? *Journal of Applied Psychology, 77*, 629-637.
- Schmit, M. J., & Ryan, A. M. (1997). Applicant withdrawal: The role of test-taking attitudes and racial differences. *Personnel Psychology, 50*, 855-867.
- Schmitt, N., Chan, D., Sacco, J. M., McFarland, L. A., & Jennings, D. (1999). Correlates of person fit and effect of person fit on test validity. *Applied Psychological Measurement, 23*, 41-53.
- Speilberger, C. D. (1980). *Test anxiety inventory manual*. California: Mind Garden.
- Zeidner, M. (1991). Test anxiety and aptitude test performance in an actual college admissions testing situation: Temporal considerations. *Personality and Individual Differences, 12*, 101-109.

Request Permission or Order Reprints Instantly

Interested in copying, sharing, or the repurposing of this article? U.S. copyright law, in most cases, directs you to first get permission from the article's rightsholder before using their content.

To lawfully obtain permission to reuse, or to order reprints of this article quickly and efficiently, click on the "Request Permission/ Order Reprints" link below and follow the instructions. For information on Fair Use limitations of U.S. copyright law, please visit [Stamford University Libraries](#), or for guidelines on Fair Use in the Classroom, please refer to [The Association of American Publishers' \(AAP\)](#).

All information and materials related to SAGE Publications are protected by the copyright laws of the United States and other countries. SAGE Publications and the SAGE logo are registered trademarks of SAGE Publications. Copyright © 2003, Sage Publications, all rights reserved. Mention of other publishers, titles or services may be registered trademarks of their respective companies. Please refer to our user help pages for more details: <http://www.sagepub.com/cc/faq/SageFAQ.htm>

[Request Permissions / Order Reprints](#)