

DISCRIMINANT AND CONCURRENT VALIDITY OF TWO COMMONLY USED MEASURES OF TEST ANXIETY¹

BRUCE A. THYER² AND JAMES D. PAPSDORF²
The University of Michigan

For each of two groups of 34 and 71 undergraduate college students respectively categorized as clinically test anxious (CTA sample) and as non-test anxious (NTA sample), evidence was sought regarding the discriminant validity and the concurrent validity of each of two measures—the 20-item self-report Test Anxiety Inventory (TAI) and the 8-item behavior analogue measure of test-taking performance entitled Anagram Solution Task. Normative information was also obtained for each measure to provide a basis for classifying students as being test anxious or as being non-test anxious. The statistical outcomes led to the conclusion that both the TAI and AST exhibit substantial discriminant validity, but relatively little concurrent validity. The creation of 90% confidence scoring intervals for each of the two measures shows considerable promise as a means for identifying students with differential levels of test anxiety. These instruments appear to afford a basis for both counseling and research purposes. Future directions in the assessment of test anxiety are discussed.

THE twofold purpose of the present paper was (a) to furnish evidence regarding the discriminant and concurrent validity of two commonly used measures of test anxiety and (b) to provide preliminary normative information for the meaningful application of the two measures. Frequently used in the assessment of test anxiety, and two measures employed in the current study were the 2-item self-report pencil-and-paper Test Anxiety Inventory (TAI) (Spielberger, 1980) and an 8-item behavior analogue measure of test-

¹ Correspondence should be addressed to the second author.

² Department of Psychology, K10 580 Union Drive, The University of Michigan, Ann Arbor, Michigan 48109.

taking performance entitled Anagram Solution Task (AST) (Sargent, 1940). Having been extensively used as a dependent variable in clinical outcome studies of test anxiety treatment programs, the TAI has a range of possible scores from 20 to 80 with higher scores being indicative of greater levels of test anxiety. The AST requires the subject sequentially to solve different anagrams under one condition intended to reflect stressful instructions and typically under a second condition involving explicit time recording.

Within the context of the use of two groups of undergraduate college subjects categorized either as non-clinically test anxious (the NTA sample) or as clinically test anxious (the CTA sample), two directional hypotheses were formulated: (a) the CTA sample would exhibit higher mean scores on the TAI than would the NTA sample and (b) the CTA sample would demonstrate a higher mean score on the AST than would the NTA sample. Data consistent with the predictions provided by these two hypotheses would be considered as evidence in support of the construct (discriminant) validity of these two measures. In addition, data in support of the concurrent validity of either one of the two measures in relation to the other serving as a criterion variable would be realized if within the CTA sample a statistically significant positive correlation occurred between TAI scores and the amount of time required to complete the AST. In other words, higher TAI scores would be associated with longer periods of time to solve the anagram tasks. Moreover, if the correlation between these two measures was positive within the CTA sample and if significantly less positive, zero, or negative in the NTA sample, additional support for the construct validity of the two measures would be realized.

Method

Subjects

Data from two subject samples were obtained. Subjects comprising the NTA sample were undergraduate students who had participated in a variety of psychological experiments and who had completed the test measures in return for partial research and course credit. The CTA sample consisted of all participants who had been enrolled in what was called the Test Anxiety Program offered by the university during the previous year. All subjects in the CTA sample had reported a history of debilitating test anxiety extending over a period of several years—*anxiety that had been elicited by anticipating or completing written or oral examinations.* Subjects in this

group had paid a \$30.00 fee to enroll in the 8-week treatment program. They had been recruited through advertisements placed in local newspapers and from posters displayed throughout the campus.

Design Features

Much of the clinical research in the area of test anxiety has been based upon the use of analogue student subjects (e.g., Bruch, 1978; Counts, Hollandsworth, and Alcorn, 1978; Holroyd, 1976; Hussian and Lawrence, 1978). The typical methodology used in these studies has been to administer some pencil-and-paper measures of test anxiety to large numbers of undergraduate students and subsequently to offer free treatment to those individuals scoring in the upper ranges of scales intended to reflect test anxiety. This form of student recruitment, which is similar to that seen in the early research conducted on the treatment of phobic disorders (Mathews, 1978), has been subject to considerable criticism on the grounds that the results obtained from research conducted on analogue subjects may not be generalizable to individuals who possess clinical levels of anxiety or fear (Barrios, 1979). The present study is one of the few available which examines the responses of clinically test anxious students and compares these responses to those of a supposedly normative sample.

Procedure

Seventy one of the subjects in the NTA group completed a battery of psychological inventories including the TAI prior to being administered the AST. In the AST, subjects were individually taken into a sound attenuated chamber where the experimenter presented them with a series of eight different anagrams taken from Sargent (1940) and previously used for research concerning the nature and treatment of test anxiety (Bruch, 1978; Denney and Rupert, 1977; Holroyd, 1978; Holroyd, Westbrook, Wolf, and Badhorn, in press; Sarason, 1961; Sarason and Palola, 1960).

The experimenter described the task to the subject and read a set of instructions adapted from Sarason (1961), which indicated that the ability to solve anagrams was related to intelligence and that the average college student should experience little trouble in solving them. The experimenter then answered any questions, presented the subject with the first anagram (printed on a 3 × 5 index card), and immediately started a stop watch placed upon the table at which the

subject and the experimenter were seated, facing one another. After the subject had solved the first anagram, the experimenter reset the stop watch, recorded the time, placed a new anagram before the subject, and started the stop watch again. This procedure was continued until the eighth and final anagram was solved. Mention should be made of the fact that this task was part of a larger series of experiments pertaining to the effects of test anxiety level, sex of subject, and distraction on cognitive performance and physiological arousal (McCann and Papsdorf, 1979; Papsdorf, Himle, Thyer, and McCann, in press).

An additional 228 of the subjects in the NTA sample completed the TAI which was administered in conjunction with other psychological inventories in a separately reported investigation of the ideational components of test anxiety (Himle, Thyer, and Papsdorf, in press). Students comprising the CTA sample ($N = 34$) similarly completed the TAI and other psychological inventories as a part of the pre-treatment assessment procedure of the Test Anxiety Program. After the measures had been completed, each subject was administered the AST as previously described. Details of the entire assessment procedures have been reported elsewhere (Thyer, Papsdorf, Himle, McCann, Caldwell, and Wickert, 1981). Copies of the assessment protocol are available from the second author.

Data Analyses

For each of the two samples, means and standard deviations were calculated on each of the two measures along with a product moment correlation coefficient between the scores on these two measures (the TAI and AST). In addition, for each measure the significance of the difference between the means of the CTA and NTA samples was determined along with the formulation of a 90 percent confidence score interval for each group. In the calculation of means, standard deviations, tests of significance of difference between means, and confidence intervals, 71 and 34 subjects were in the NTA and CTA samples, respectively. For the correlation coefficient between the TAI and AST measures, 71 subjects were in the NTA sample, but only 19 in the CTA sample.

Results

The following major statistical outcomes are summarized:

1. In the instance of the TAI measure, the mean and standard deviation for the CTA sample were 61.26 and 10.76, respectively;

for the NTA sample, 40.18 and 1.92. The positive mean difference of 21.06 in a one-tailed test was significant beyond the .0005 level.

2. Relative to the AST measure, the mean and standard deviation for the CTA group were 216.64 and 89.58, respectively; for the NTA sample, 177.71 and 73.45. The positive mean difference of 38.93 in a one-tailed test was significant beyond the .0005 level. (Score units were in seconds.)

3. For the CTA sample, a correlation of $-.28$ occurred between the TAI and AST measures, which in terms of a one-tailed test with a hypothesized positive correlation was not statistically significant. On the other hand, for the NTA sample the correlation $.35$ in the circumstance of a non-directional hypothesis was statistically significant beyond the $.01$ level.

4. Relative to the TAI measure, the 90 percent confidence interval in scores of the NTA sample extended from 39.05 to 41.33; correspondingly, in scores of the CTA sample, from 58.14 to 64.39. With respect to the AST measure, the 90 percent interval in scores for the NTA sample ranged from 163.18 to 192.24 (seconds); for the CTA sample, from 190.64 to 242.62 (seconds).

Conclusions

The following conclusions were apparent from the statistical outcomes:

1. In light of the significance of the difference between means of the CTA and NTA samples in the direction hypothesized on each of the two measures, it appears that positive support exists for their discriminant validity.

2. That the correlation of $-.28$ between the TAI and AST measures was opposite in direction to that anticipated would argue against the concurrent validity of either measure relative to the other one as a criterion variable and very possibly against the construct validity of either measure within the CTA sample. However, in the NTA sample, which was probably representative of most college students without a debilitating level of test anxiety, the positive correlation of $.35$ would indicate a modest relationship between a measure of the anxiety construct portrayed in the TAI measure and the measure of an associated construct of inferred stress or task anxiety reflected by the AST. This correlation would suggest a small degree of concurrent validity for each measure in the context of the other serving as a criterion variable for students judged not to show highly deviant or almost pathological levels of test anxiety.

3. The virtual lack of overlap in the 90% confidence score

intervals between the CTA and NTA samples relative to the AST measure and the complete lack of overlap with respect to the TAI measure suggest that useful normative information has been obtained for diagnostic and classification purposes that could be employed to substantial advantage by counselors in a university setting as well as by psychologists in clinical research.

Discussion

The TAI does seem to possess sufficiently high discriminant validity to justify its continued use as a screening instrument for research purposes and as a dependent variable in clinical outcome studies. The normative TAI value for the NTA sample was very close to a score of 40, whereas a score greater than 50 in most instances was evident for the CTA sample. Certainly the lack of overlap in the confidence score intervals suggested substantial promise for the TAI in differentiating college students with contrasting levels of test anxiety.

Performance on the AST was also observed to differentiate between those who fell in the two samples. Although the AST is probably not so efficient a method for initial screening and classification purposes in test anxiety treatment programs as is the TAI, it may serve as a useful dependent variable in clinical outcome studies. In a recent experiment conducted at the same laboratory as the one employed in this investigation, post-treatment mean anagram solution times were found to improve (decrease) substantially relative to pre-treatment solution times for a group of subjects participating in an experiment on the efficacy of a cognitive behavior therapy procedure (Thyer et al., 1981).

As an analogue of test-taking performance, the AST is probably less subject to artifact effects (e.g., placebo factors, experimental demands, or client expectations) than are traditional measures of test anxiety. Accordingly, the AST has much to recommend it for behavioral assessment purposes in the field of test anxiety study. The AST measures a frequently untapped response channel—namely that of cognitive functioning under conditions probably reflecting mild evaluative stress. The measure possesses promising discriminant validity, and initial normative data are now available for preliminary classification of students who manifest behaviors judged to be clinically test anxious or non-test anxious.

The small to nonexistent relationship obtained between scores on the TAI and AST measures (particularly in the CTA sample) was not altogether unexpected in view of the small sample size and in light of

the presence of quite low correlation among the three response channels of self-report information, visual-motoric behavior, and physiological arousal. It is also possible that these two measures could be validly portraying two or more separate but nonlinearly related aspects of test anxiety such as a worry or cognitive component and an emotional or autonomic arousal element or even a test-taking impairment factor. Obviously research is needed to develop instruments to identify differentially which aspect of test anxiety is pertinent to the treatment of a given individual exhibiting behavior patterns indicative of debilitating levels of anxiety in responding to examination tasks. Further research is now underway in the laboratory at which this study was done to develop assessment procedures cutting across multiple response channels for the identification of the individual who reports great stress and anxiety in test-taking experiences. It is hoped that behavioral assessment data with prognostic significance and with treatment validity will be forthcoming.

REFERENCES

- Barrios, B. Publication trends on behavior therapy analogue research on phobias. *Journal of Behavior Therapy and Experimental Psychiatry*, 1979, 10, 203-205.
- Bruch, M. Type of cognitive modeling, Imitation of modeled tactics, and modification of test anxiety. *Cognitive Therapy and Research*, 1978, 2, 147-164.
- Counts, D., Hollandsworth, J., and Alcorn, J. Use of EMG biofeedback and cue controlled relaxation in the treatment of test anxiety. *Journal of Consulting and Clinical Psychology*, 1978, 2, 147-164.
- Denney, D. and Rupert, P. Desensitization and self control in the treatment of test anxiety. *Journal of Counseling Psychology*, 1977, 24, 272-280.
- Himle, D., Thyer, B., and Papsdorf, J. Relationships between rational beliefs and anxiety. *Cognitive Therapy and Research* (in press).
- Holroyd, K. Cognition and desensitization in the group treatment of test anxiety. *Journal of Consulting and Clinical Psychology*, 1976, 44, 991-1001.
- Holroyd, K. Effectiveness of an 'attribution therapy' manipulation with test anxiety. *Behavior Therapy*, 1978, 9, 526-534.
- Holroyd, K., Westbrook, T., Wolf, M., and Badhorn, E. Performance, cognition and physiological responding in test anxiety. *Journal of Abnormal Psychology* (in press).
- Hussian, R. and Lawrence, P. The reduction of test, state and trait anxiety by test specific and generalized stress inoculation training. *Cognitive Therapy and Research*, 1978, 2, 25-37.
- McCann, B. and Papsdorf, J. *Bilateral hand temperature, test*

- anxiety and anagram solution stress*. Paper presented at the annual meeting of the Biofeedback Society of America, 1979.
- Mathews, A. Fear reduction research and clinical phobias. *Psychological Bulletin*, 1978, 85, 390-404.
- Papsdorf, J., Himle, D., Thyer, B., and McCann, B. Anagram solution time in the presence of a distracting experimenter; sex differences and test anxiety level. *Perceptual and Motor Skills* (in press).
- Sarason, I. The effects of anxiety and threat on the solution of a difficult task. *Journal of Abnormal and Social Psychology*, 1961, 62, 165-168.
- Sarason, I. and Palola, E. The relationship of test and general anxiety, difficulty of task and experimental instruction to performance. *Journal of Experimental Psychology*, 1960, 59, 185-191.
- Sargent, S. Thinking processes at various levels of difficulty. *Archives of Psychology*, 1940, No. 249.
- Spielberger, C. *The Test Anxiety Inventory* (TAI). Palo Alto, Calif.: Consulting Psychologists Press, 1980.
- Thyer, B., Papsdorf, J., Himle, D., McCann, B., Caldwell, S. and Wickert, M. In-vivo distraction coping in the treatment of test anxiety. *Journal of Clinical Psychology*, 1981, 37, 754-764.