Journal of Anxiety Disorders, Vol. 6, pp. 55–61, 1992
Printed in the USA.

0887-6185/92 $5.00 + .00

Reliability and Validity of the Beck Anxiety Inventory

THOMAS FYDRICH, PH.D.

Temple University Medical School and Agoraphobia and Anxiety Treatment Center

DEBORAH DOWDALL, B.A.

Agoraphobia and Anxiety Treatment Center, Bala Cynwyd, PA

DIANNE L. CHAMBLESS, PH.D.

The American University

Abstract — Two studies were conducted to further psychometric research on the recently developed Beck Anxiety Inventory (BAI). In Study 1 the test-retest reliability and internal consistency of the scale were examined with a sample of 40 outpatients having anxiety disorders. The BAI proved highly internally consistent (Cronbach's alpha = .94) and acceptably reliable over an average time lapse of 11 days (r = .67). Study 2 was conducted to assess the convergent and discriminant validity of the BAI vis à vis anxiety and depression and in comparison to the widely used Trait Anxiety measure from the State-Trait Anxiety Inventory. Seventy-one outpatients with anxiety disorders completed the revised State-Trait Anxiety Inventory, the Beck Depression Inventory, and daily diary ratings of anxiety and depression in addition to the BAI. The BAI fared better on tests of convergent and discriminant validity than did Trait Anxiety. The correlation between the BAI and Diary Anxiety was significantly higher than that between BAI and Diary Depression, and, compared to Trait Anxiety, the BAI was significantly less confounded with depression as measured by the BDI. Scores for STAI-Y Trait Anxiety were highly confounded with measures of depression, but results for the STAI-Y State scale were more positive.

Symptoms of anxiety and depression appear to be closely related. Indeed, some authors (e.g., Barlow et al., 1986) consider the overlap so extensive that a useful distinction cannot be made. In particular, this overlap manifests itself

---

Thomas Fydrich is now at the University of Heidelberg. The authors wish to thank the therapists at the Agoraphobia and Anxiety Treatment Center for their cooperation in the data collection process and Edward Gracely, Ph.D., for statistical consultation. Requests for reprints should be sent to Dianne L. Chambless, Ph.D., Department of Psychology, The American University, Washington, D.C. 20016-8062, U.S.A.

in self-report measures of anxiety and depression. Dobson (1985) reviewed 34 studies reporting correlations between self-report measures of anxiety and depression, and found a mean correlation of .61. From an analysis of the psychometric properties of several widely used measures of anxiety and depression, Tanaka-Matsumi and Kameoka (1986) concluded that measures of anxiety and depression were as highly intercorrelated with one another (mean $r = .64$) as were anxiety measures among themselves (mean $r = .67$). These authors even observed paradoxical patterns wherein the STAI-State scale correlated higher with two measures of depression (Zung-D and BDI; $r = .74$ and $r = .73$, respectively) than with other measures of anxiety (STAI-Trait, Taylor Manifest Anxiety Scale, and Zung-A; .66 and .53, and .60, respectively).

Beyond true overlap in the clinical picture of the disorders, one factor contributing to the lack of discrimination on self-report measures is the similarity in items incorporated in these scales. For example, an item on the 1970 version of the Trait form of the STAI (Spielberger et al., 1970) reads, "I feel blue." Second, some of the most widely used anxiety scales were initially developed largely on student populations. As such, items more typical of clinical anxiety states may be under-represented. The lack of cross-validation of these scales with clinical samples is troubling.

In this context, the development of a new self-report scale for anxiety, the Beck Anxiety Inventory (BAI: Beck & Steer, 1990; Beck, Epstein, Brown, & Steer, 1988), is especially welcome. The BAI was developed on, and for use with, psychiatric patients; items were selected with a strong emphasis on discriminating anxiety from depression. Although the BAI was, as is appropriate, modestly correlated with the revised Hamilton Depression Inventory ($r = .25$), it was more highly and substantially correlated with the revised Hamilton Anxiety Inventory ($r = .51$), and it successfully discriminated anxious from depressed patients.

The present studies were designed to provide additional empirical data on the reliability and validity of the BAI with clinical samples of anxiety patients. The convergent and discriminant validity of the BAI are contrasted with those of a widely used measure, the State-Trait Anxiety Inventory.

STUDY 1

Subjects

Subjects were 40 outpatients (21 women, 19 men) with a primary diagnosis of anxiety disorder who had applied for treatment at the Agoraphobia and Anxiety Treatment Center. Diagnoses were determined according to DSM-III-R criteria (American Psychiatric Association, 1987) on the basis of a comprehensive intake evaluation by a licensed psychologist specializing in anxiety disorders. Axis I diagnoses were as follows: 21 had panic disorder with agoraphobia, 9 panic disorder without agoraphobia, 5 simple phobia, 3 obsessive-compulsive disorder, 1 generalized anxiety disorder, and 1 had an anxiety disorder not otherwise specified.

Subjects' mean age was 34 years ($SD = 8.4$), and the mean duration of anxiety symptoms was 11.8 years ($SD = 12.6$). Socioeconomic status ($M = 32.9$,

SD = 16.6) was derived according to the Hollingshead Two-Factor Index of Social Position (Hollingshead, 1957) with a mode of Social Class III, indicative of a middle-class population.

Measure

The BAI consists of 21 items, most of which closely represent DSM-III-R criteria for panic disorder. Clients rate items according to how much they are bothered by the particular symptom; each item is on a four-point scale ranging from 0 (not at all) to 3 (severely, I could barely stand it). Thirteen items describe physical or physiological symptoms (e.g., heart pounding), five represent clearly cognitive aspects of anxiety (fear of the worst), and three have a physical as well as cognitive connotation (e.g., terrified). Beck, Epstein et al. (1988) reported excellent internal consistency (Cronbach's alpha = .92) and a one-week retest reliability coefficient of .75.

Procedure

The BAI was completed by clients before the initial evaluation and readministered prior to the first treatment session.

STUDY 2

In this study, we conducted a multitrait-multimethod correlational analysis of the convergent and discriminant validity of the BAI. The two methods were daily diary ratings of mood v. questionnaire, and the two traits were depression v. anxiety. In psychometric research it is desirable to examine not only how well a new scale fares in an absolute sense, but also how its performance compares to that of other measures already available. Accordingly, we included the widely used State-Trait Anxiety Inventory as a bench mark.

Subjects

The sample comprised 71 clients (47 women, 24 men) with a primary diagnosis of anxiety disorder in treatment at the Agoraphobia and Anxiety Treatment Center. Forty-nine had panic disorder with agoraphobia, 11 panic disorder without agoraphobia, 7 social phobia, 2 obsessive-compulsive disorder, 1 generalized anxiety disorder, and 1 had an anxiety disorder not otherwise specified. The clients averaged 35.8 years in age (*SD* = 11.1), and were most commonly middle class (mode = Hollingshead Social Class III). The mean duration of anxiety symptoms was 15.0 years (*SD* = 14.0).

Measures

The State-Trait Anxiety Inventory (STAI, Form-Y: Spielberger, 1983) is a widely used self-report measure for state and trait anxiety. According to its manual, the revised form (Form-Y) is a purer measure of anxiety with better discrimination between symptoms of anxiety and depression than the previous

1970 version. Spielberger (1983) reports a mean KR-20 coefficient > .90 for both the state and the trait parts of the revised STAI. For a sample of high-school students, 30-day test-retest reliability for the trait scale was .73.

The Beck Depression Inventory (BDI: Beck & Steer, 1987) is a widely used self-report measure of depression. The BDI has been administered to a variety of clinical populations, and its reliability and validity have been extensively documented (Beck, Steer, & Garbin, 1988).

The Weekly Record of Anxiety and Depression (Barlow & Cerny, 1988) is a self-report monitoring form. Subjects are asked to rate their levels of anxiety and depression on a scale from 0 (none) to 8 (as much as you can imagine) daily for a one-week period. Subjects were instructed to complete their forms at the same time each day, and the daily ratings were averaged to yield the weekly Diary Depression or Diary Anxiety score. Diary ratings were included to provide a different method of assessing depression and anxiety than that of the global self-report questionnaire. We reasoned that the increased specificity occasioned by daily ratings would enable clients to discriminate more clearly between anxious and depressed mood and would yield a more genuine picture of the client's mood during the entire week (which the BAI and BDI purport to assess) than an overall self-rating made at the week's end. The latter would be more likely to be affected by state mood effects at the time of completion.

Procedure

Clients were instructed to complete the Weekly Record of Anxiety and Depression for a period of seven days. Within one day after the completion of these records, clients were given the BAI, BDI, and the STAI (Form-Y) in randomized order. The various questionnaires require that the client respond using different time frames. The reference point is right now for the STAI-State, the past week for the BAI and BDI, and in general for the STAI-Trait. Accordingly, somewhat lower correlations were to be expected between the Diary Anxiety (the average of the previous week's ratings) and the two STAI measures than between the BAI and these ratings.

GENERAL RESULTS AND DISCUSSION

For the Study 1 sample, who took the BAI at intake, a broad range of initial BAI scores was obtained (1–49). The scores were normally distributed (Kolmogeroff-Smirnoff test $p = .94$) around a mean of 23.9 ($SD = 13.7$). An ANOVA [$F(2, 37) = 2.97$, $p = .06$] was used to compare the scores of three diagnostic subgroups in this sample: panic disorder with agoraphobia ($M = 28.29$, $SD = 14.29$), panic disorder without agoraphobia ($M = 22.22$, $SD = 12.21$), and other anxiety disorders ($M = 16.30$, $SD = 10.48$). The two groups of clients with panic scored significantly higher than clients with other anxiety disorders ($p = .04$).

Reliability

To assess internal reliability, Cronbach's alpha was computed for responses to the first administration of the BAI. Our data replicated those of Beck,

Epstein et al. (1988). The BAI was highly internally consistent, alpha = .94. The Pearson test-retest correlation for an average time lapse of 11 days was .67. Although this is somewhat lower than the reliability of .75 reported by Beck et al., the difference would appear to be due to our longer test-retest interval. If we restrict our sample to those clients who were retested in seven days (the interval used in Beck et al.'s study), we obtain a very similar coefficient of .73 ($n = 26$). These analyses show that the scale has high internal consistency and sufficient retest reliability for a test that, by definition and instruction, does not aim to diagnose trait anxiety.

Validity

A multi-trait, multimethod correlation matrix was computed and is reported in Table 1. Correlations of the BAI and STAI with Diary Anxiety represent monotrait-heteromethod coefficients, while correlations of the BAI and STAI with the BDI represent the heterotrait-monomethod coefficients.

The BAI test-retest reliability coefficient of .67 and the STAI-Trait coefficient of .73 set the expected upper limits for correlation of these scales with others. Accordingly, the correlation of .54 obtained between the BAI and Diary Anxiety is a very robust convergent validity coefficient. This correlation is significantly higher ($p < .05$) than that of the BAI and Diary Depression by Olkin's test (Steiger, 1980), thereby supporting the discriminant validity of the scale. Reflecting shared method variance as well as the overlap of anxiety and depression, the BAI does have a substantial correlation with the BDI, virtually the same found by Beck et al. However, this correlation is significantly lower than that of the BDI and the Trait Anxiety scale (Olkin's test $p < .01$). Although we used diary ratings to assess concurrent validity, whereas Beck et al. used an interview, the effect sizes for convergent validity in the two studies are quite similar. Such convergence across disparate assessment methods provides particularly encouraging support for the construct validity of the BAI.

On the other hand, Trait Anxiety not only failed to correlate significantly with Diary Anxiety but also often correlated more strongly with measures of depression than with measures of anxiety. Because the Trait scale asks clients to rate how they feel in general, we would expect that it would have lower

TABLE 1
MULTITRAIT-MULTIMETHOD CORRELATION MATRIX OF ANXIETY AND DEPRESSION MEASURES FOR 71 ANXIOUS OUTPATIENTS

	BAI	Trait	State	Diary Anxiety	BDI
STAI-Trait	0.58***				
STAI-State	0.47**	0.75***			
Diary Anx.	0.54***	0.34	0.53***		
BDI	0.50***	0.73*	0.59***	0.37*	
Diary Dep.	0.38*	0.36	0.40*	0.62***	0.42**

Bonferroni-corrected probabilities: *$p < .05$, **$p < .01$, ***$p < .001$.

correlations with weekly diary ratings than the more comparably framed BAI. However, the correlation of Trait Anxiety with the BDI is so high (.73) that it matches the test-retest reliability of the scale. Thus, even in its revised form, STAI-Trait has as much variance in common with self-reported depression as with the closely related concept of state anxiety.

The results of analyses on State Anxiety were more supportive of that scale's validity. Although State Anxiety did not discriminate as well as the BAI, its correlation with Diary Anxiety tended to be higher than that with Diary Depression (Olkin's test, $p = .075$), and it was less highly correlated with the BDI than was the Trait scale.

Several limitations to the present study should be considered in interpreting these results. First, although it has proved useful in prior research, the Weekly Diary Form has not itself been subjected to formal psychometric evaluation. The present results indicate considerable overlap between Diary Anxiety and Depression ratings (.62). This high correlation makes the degree of discrimination achieved by the BAI all the more remarkable. Second, the correlations of the anxiety questionnaires, particularly the BAI and State Anxiety, with Diary Anxiety may be somewhat exaggerated by the data collection procedure. Having rated their anxiety daily for the previous week may have heightened clients' awareness of their anxious symptoms and increased the precision of their ratings.

Finally, the present results may be more supportive of the BAI than would data collected on a different sample of anxious patients. Clients with panic disorder predominated in our sample and in that of Beck, Epstein et al. (1988). Examination of the BAI's items suggests that the BAI may be most sensitive to the symptoms of such patients: The items are clearly dominated by physical symptomatology. In light of the importance of cognitive aspects of anxiety, particularly for generalized anxiety disorder and obsessive-compulsive disorder, the content validity of the scale with different anxious patients might be improved by a broader range of items. In our sample, clients with panic disorder scored higher on the BAI than those with other anxiety disorders. Given the composition of the scale, it is impossible to tell whether they are genuinely more anxious or whether the scale items are simply more representative of panic-related anxiety.

CONCLUSIONS

The results from the present study confirm those of Beck, Epstein et al. (1988). The BAI is internally consistent and reliable and appears to have better convergent and discriminant validity than the most widely used self-report measure of anxiety. However, all of the anxiety measures we included were highly intertwined with measures of depression. These results support the notion of a general mood factor to which clients refer when they rate anxiety and depression. Thus, it seems likely that self-report measures differentiate less between anxiety and depression than standardized clinician-rated measures, such as the revised Hamilton Depression and Anxiety scales (Riskind et al., 1987) and diagnostic schedules such as the Structured Clinical Interview for DSM-III-R (Spitzer et al., 1987).

REFERENCES

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders.* Washington: Author.

Barlow, D. H., DiNardo, P. A., Vermilyea , B. B., Vermilyea, J. A., & Blanchard, E. B. (1986). Co-morbidity and depression among the anxiety disorders: Issues in diagnosis and classification. *Journal of Nervous and Mental Disease*, **174**, 63–72.

Barlow, D. H., & Cerny, J. A. (1988). *Psychological treatment of panic.* New York: Guilford.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, **56**, 893–897.

Beck, A. T., & Steer, R. A. (1987). *Manual for the revised Beck Depression Inventory.* San Antonio, TX: Psychological Corporation.

Beck, A. T., & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, **8**, 77–100.

Dobson, K. S. (1985). The relationship between anxiety and depression. *Clinical Psychology Review*, **5**, 307–324.

Hollingshead, A. B. (1957). *Two factor index of social position.* (Available from A.B. Hollingshead, Yale University, 1965 Yale Station, New Haven, CT 06520.)

Riskind, J. R., Beck, A. T., Brown, G., & Steer, R. A. (1987). Taking the measure of anxiety and depression: Validity of the reconstructed Hamilton scales. *Journal of Nervous and Mental Disease*, **175**, 474–479.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory* (Form Y). Palo Alto: Consulting Psychologists Press.

Spielberger, C., Gorsuch, A., & Lushene, R. (1970). *The State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Spitzer, R. L., Williams, J. B. W., & Gibbon, M. (1987). *Structured Clinical Interview for DSM-III-R — Outpatient Version.* New York: Biometrics Research Department, New York State Psychiatric Institute.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, **87**, 245–251.

Tanaka-Matsumi, J., & Kameoka, V. A. (1986). Reliabilities and concurrent validities of popular self-report measures of depression, anxiety, and social desirability. *Journal of Consulting and Clinical Psychology*, **54**, 328–333.