Journal of Consulting and Clinical Psychology
1993, Vol. 61, No. 4, 631-638

0022-006X/93/$3.00

Distinguishing Depression and Anxiety in Self-Report: Evidence From Confirmatory Factor Analysis on Nonclinical and Clinical Samples

Lisa A. Feldman

Psychologists believe that anxiety and depression self-report scales tap distinct constructs. This assumption was tested by using confirmatory factor analysis on mood data from nonclinical samples (K. S. Dobson, 1985a; I. H. Gotlib, 1984; J. Tanaka-Matsumi & V. A. Kameoka, 1986) and a clinical sample (J. Mendels, N. Weinstein, & C. Cochrane, 1972). These analyses provide evidence that anxiety and depression self-report scales do not measure discriminant mood constructs and may therefore be better thought of as measures of general negative mood rather than as measures of anxiety and depression per se.

Most people consider anxiety and depression to be distinct mood experiences. People conceptually discriminate between anxiety and depression when they define each mood experience (Davitz, 1969; Pennebaker, 1982) and when they assess the similarity between mood terms (e.g., Bush, 1973; Russell, 1980). The psychological theories of clinicians and researchers reflect the intuitions and language of laypeople (e.g., Beck, 1976; Davitz, 1969; Higgins, 1987; Izard, 1977; Klerman, 1977; Ortony, Clore, & Collins, 1988; Pennebaker, 1982; Wallbott & Scherer, 1989). Psychologists describe depression as the experience of being sad, blue, and gloomy; depression is usually associated with the perception of loss, events that occurred in the past, and decreased autonomic activity. Alternatively, psychologists describe anxiety as feelings of fear, dread, and tension; anxiety is typically associated with perceptions of threat or uncertainty, anticipation of future events, and increased autonomic activity.

Despite the firmly held convictions of laypeople and theorists that anxiety and depression are distinct mood experiences, researchers have generally been unable to separate them reliably in self-report. Evidence that respondents' reports of anxiety and depression are indistinguishable comes from four areas of research. First, many researchers have found such large correlations between various self-report measures of anxiety and depression that, given the reliability of the measures, no meaningful discrimination between self-reported depression and anxiety can be identified. (e.g., L. A. Clark & Watson, 1991; Dobson, 1985a, 1985b; Tanaka-Matsumi & Kameoka, 1986). Second, the strong correlations between the anxiety and depression self-report measures have left investigators unable to extract distinct mood groups when attempting to examine anxiety and depression as orthogonal dimensions in a 2 × 2 factorial design (see Gotlib, 1984, for a detailed review). Third, most exploratory factor analyses of self-reported anxiety and depression ratings have produced a general negative mood or dysphoria factor rather than distinct depression and anxiety factors (e.g., Gotlib, 1984; Gotlib & Meyer, 1986; Zuckerman, Lubin, & Rinck, 1983). Fourth, mood induction studies indicate that attempts to elicit depression also increase self-reported anxiety. Similarly, efforts to induce fear also produce elevations in self-reported depression (see Polivy, 1981, for a review).

Furthermore, the pattern of findings for self-reported anxiety and depression in the clinical literature parallels the findings described for studies of nonclinical and subclinical samples. In general, clear and consistent differentiation between measures of anxiety and depression is rarely achieved from the self-reports of patients (L. A. Clark, 1989; L. A. Clark & Watson, 1991). The clinical findings are characterized by high correlations between self-report measures of anxiety and depression, the extraction of a single factor with significant loadings from both anxiety and depression measures, high comorbidity rates for diagnoses of the two disorders, and nonspecific drug responses in both disorders (for a review, see D. A. Clark, Beck, & Stewart, 1990; L. A. Clark, 1989; Maser & Cloninger, 1990). Discrimination between measures of anxiety and depression has not been obtained in heterogeneous groups of patients (Mendels, Weinstein, & Cochrane, 1972) or in distinct diagnostic groups (e.g., Downing & Rickels, 1974; Lipman, 1982; Prusoff & Klerman, 1974). Even measures that have been designed for discriminant validity are moderately correlated with each other (Beck, Epstein, Brown, & Steer, 1988; D. A. Clark et al., 1990).

Is it the case that anxiety and depression are empirically indistinguishable as far as self-report methods allow? Is there any discriminant validity of depression and anxiety measures (cf. Campbell & Fiske, 1959)? Some may argue that the strong relationship between anxiety and depression measures reflects a simple co-occurrence between two phenomenologically independent mood states. There is evidence to suggest, however, that both moods share a strong general negative mood compo-

Lisa A. Feldman, University of Waterloo, Waterloo, Ontario, Canada.

This research was conducted as part of my doctoral dissertation and was supported in part by a doctoral fellowship from the Social Sciences and Humanities Research Council of Canada.

I thank Geoffrey Fong, Michael Ross, Erik Woody, Ken Bollen, Steven Breckler, Michael Newcomb, Deborah Fitzpatrick, and Jonathan Golden for their very helpful comments on earlier versions of the manuscript.

Correspondence concerning this article should be addressed to Lisa A. Feldman, who is now at Department of Psychology, 417 Moore Building, Pennsylvania State University, University Park, Pennsylvania 16802.

nent (L. A. Clark & Watson, 1991). Perhaps self-report measures of anxiety and depression are more accurately considered as measures of general negative mood.

Confirmatory factor analysis provides the most powerful method for determining whether anxiety and depression can be meaningfully separated by current measures. Confirmatory factor analysis has two advantages over exploratory factor-analytic procedures. In a test of a hypothesized structure's or model's consistency with the observed covariation in the data, confirmatory factor analysis imposes theoretically substantive constraints on the solution. In contrast, in exploratory factor analysis, the solution is determined by arbitrary statistical constraints. Furthermore, models resulting from competing hypotheses can be compared directly with one another using confirmatory procedures; this is not possible using exploratory procedures.

In this investigation, I conducted a series of confirmatory factor analyses on several sets of published correlation matrices to determine directly whether anxiety and depression could be distinguished with self-report measures. A two-factor anxiety and depression model was compared with a more general one-factor negative mood model to determine whether anxiety and depression can be meaningfully separated in nonclinical and in clinical samples.

Because much of the mood research to date has been conducted on nonclinical and subclinical samples, it is particularly important to demonstrate discrimination in anxiety and depression self-report measures when used with analogue respondents. I reanalyzed the correlation matrices of anxiety and depression measures published by Dobson (1985a), Gotlib (1984), and Tanaka-Matsumi and Kameoka (1986) to test whether a two-factor anxiety and depression model fit the data better than a one-factor negative mood model. I selected the data from these studies for reanalysis because they included the most commonly used self-report measures of anxiety and depression. I reanalyzed the correlation matrix of anxiety and depression measures from Mendels et al. (1972) to examine the relationship between anxiety and depression in a clinical population. I selected the Mendels et al. (1972) data for reanalysis because, to my knowledge, it was the only study that includes enough self-report measures or enough participants to conduct a confirmatory factor analysis.

Method

Studies Selected

Dobson (1985a). In Dobson's study, nine self-report measures of anxiety and depression were administered to 71 female undergraduate students. The depression scales used in the present analyses[1] were the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), the Depression subscale of the Multiple Affect Adjective Check List (MAACL–D; Zuckerman & Lubin, 1965), and the Costello–Comrey Depression Scale (CC–D; Costello & Comrey, 1967). The anxiety scales used were the State–Trait Anxiety Inventory (STAI), Trait version (Spielberger, Gorsuch, & Lushene, 1970), the Anxiety subscale of the Multiple Affect Adjective Check List (MAACL–A; Zuckerman & Lubin, 1965), and the Costello–Comrey Anxiety Scale (CC–A; Costello & Comrey, 1967).

Gotlib (1984). Anxiety and depression self-report scales, along with other measures of psychopathology and assertiveness, were administered to 147 male and 296 female undergraduate students. The depression scales used in the present analyses were the BDI, the MAACL–D, and the Depression subscale of the Symptom Checklist—90–Revised (SCL–D; Derogatis, Lipman, & Covi, 1973). The anxiety scales used were the STAI, the MAACL–A, and the Anxiety subscale of the Symptom Checklist—90–Revised (SCL–A; Derogatis et al., 1973).

Tanaka-Matsumi and Kameoka (1986). Anxiety, depression, and social desirability self-report scales were administered to 135 male and 256 female undergraduate students. The depression scales used in the present analyses were the Zung Depression Scale (ZUNGD; Zung, 1965), the BDI, and the Lubin Depression Adjective Checklist (DACL; Lubin, 1967). The anxiety scales used were the STAI (both State and Trait versions) and the Zung Self-Rating Anxiety Scale (ZUNGA; Zung, 1971).

Mendels et al. (1972). Anxiety and depression self-report scales, along with the Minnesota Multiphasic Personality Inventory, were administered to 76 White women, ranging from 21 to 65 years of age, who were hospitalized on an acute psychiatric inpatient service. The depression scales used were the BDI, the MAACL–D, and the CC–D. The anxiety scales used were the MAACL–A and the CC–A.

Procedure

The correlation matrices were subjected to confirmatory factor analyses using LISREL VI (Jöreskog & Sörbom, 1984).[2] A two-factor anxiety and depression model was compared with a more general one-factor negative mood model for each data set. In the two-factor model, anxiety measures acted as indicators of an anxiety latent construct, and depression measures acted as indicators of a separate depression latent construct. The parameter representing the correlation between the latent anxiety and depression constructs (ϕ) was left free to be estimated. In the one-factor model, anxiety and depression measures were indicators for a single negative mood factor. The parameter ϕ was set to equal 1.0, indicating that only one latent construct was represented in these analyses. In all analyses, errors were correlated between anxiety and depression subscales taken from the same measure to control for the method variance shared by subscales of the same instrument.

Statistical Assumptions

Multivariate normality. Maximum likelihood estimation of the parameters in a confirmatory factor analysis requires the assumption that the observed variables have a multivariate normal distribution in the population. Large samples are typically required to justify the use of estimation procedures that require the assumption of multivariate normality in the population distributions (Breckler, 1990); however, precise guidelines for adequate sample sizes have not been given, and researchers have tended to disagree on the smallest sample size that they would find acceptable (Anderson & Gerbing, 1984; Breckler,

[1] Several measures of anxiety and depression included in the correlation matrices were not included in the present analyses either because sample size did not permit the inclusion of all of the measures or because the measures used a categorical (true–false) response scale. The average correlations of each scale with anxiety measures and depression measures in every data set were not different for those scales included in the analyses and for those scales that were not included.

[2] Although some researchers have recommended the use of covariance rather than correlation matrices in confirmatory factor analyses (Hull, 1989), the models evaluated in the present study were scale invariant (Cudeck, 1989) and thus could be evaluated using correlation matrices.

1990; Gerbing & Anderson, 1985; Hayduk, 1987). Some researchers have suggested that maximum likelihood estimation procedures should only be attempted when the ratio of the number of variables to the number of respondents, as well as the ratio of the number of parameters to be estimated to the number of respondents, is above 1:4 (Tanaka, 1987) or 1:5 (Bentler & Chou, 1988, as cited in Hull, 1989). Both the ratio of the number of variables to the number of participants and the ratio of the number of parameters to be estimated to the number of participants were adequate for the analyses of all four data sets.

Ordinal data. Confirmatory factor analysis requires that the variables being measured by the indicators are continuous in the population. The measures under evaluation in the present analyses, however, use scales that are better classified as ordinal (e.g., 4-point scales). This was not problematic in the present analyses, because the indicators being factored were scale scores rather than items (Bernstein & Teng, 1989).

Results

The four sets of confirmatory factor analyses provided fairly consistent evidence that anxiety and depression self-report scales measure a single mood construct. Several fit indices indicated that the one-factor models fit the data as well as the two-factor models do for most data sets. The parameter estimates of the confirmatory factor analyses are presented in Table 1.

Chi-Square Tests

The chi-square test evaluates whether the covariance matrix generated by the theoretical model deviates significantly from the observed covariance matrix. Some researchers have called the chi-square test a "lack-of-fit" index (Mulaik et al., 1989). A well-fitting model is associated with a nonsignificant chi-square, indicating that the discrepancy between the observed covariance matrix and the matrix generated by the theoretical model is not significant. The chi-square statistics for the four sets of analyses are presented in Table 2. Chi-square statistics were significant for all of the 4 one-factor models and 3 of the 4 two-factor models, suggesting some degree of lack of fit for the majority of the models tested.

The only difference between the one- and two-factor models is that in the latter the correlation between the factors was estimated. It was possible to determine whether the two-factor models fit the data better than the one-factor models do by testing the difference between the chi-square statistics. The difference chi-square statistics were not significant for two of the nonclinical samples, which suggests that the one- and two-factor models fit these two data sets equally (see Table 2). The change in chi-square value was significant for the Gotlib (1984) and the Mendels et al. (1972) data sets, suggesting the possibility that the two-factor model fit the data better than the one-factor model in these samples.

Fit Indices

Researchers have been dissatisfied with the chi-square test as a measure of fit. It does not provide an adequate assessment of the fit of a model because of its dependence on sample size and because the statistic does not provide an indication of degree of fit (see Bentler, 1990; Mulaik et al., 1989). These dissatisfactions with the chi-square statistic as a measure of fit have led researchers to develop better fit measures.

Comparative fit index (CFI). A normed-fit index, the CFI (see Bentler, 1990; Bentler & Bonnett, 1980), evaluates the adequacy of the specified model in relation to a baseline model. The most restricted baseline model (called a null model) was used in the calculation of the CFI; a null model hypothesizes no relationship among the observed variables (i.e., every measure is an indicator for a separate latent variable). Fit coefficients range from 0 to 1, with higher values indicating greater fit. Bentler and Bonnett (1980) have suggested that a normed-fit index of .90 or greater indicates that the hypothetical model fits

Table 1
Parameter Estimates Associated With the One- and Two-Factor Models

	1-Factor model		2-Factor model	
Study/measure	ϕ	λ	ϕ	λ
Dobson (1985a)	1.0		.96	
BDI		.88		.89
MAACL–D		.71		.72
CC–D		.86		.87
A-Trait		.90		.92
MAACL–A		.77		.78
CC–A		.63		.64
Gotlib (1984)	1.0		.92	
BDI		.79		.81
MAACL–D		.70		.71
SCL–D		.81		.82
A-Trait		.83		.87
MAACL–A		.73		.76
SCL–A		.70		.69
Tanaka-Matsumi and Kameoka (1986)	1.0		.999	
ZUNGD		.81		.81
BDI		.82		.82
DACL		.71		.71
ZUNGA		.77		.77
A-State		.78		.78
A-Trait		.89		.89
Mendels, Weinstein, and Cochrane (1972)	1.0		.82	
BDI		.86		.87
MAACL–D		.69		.69
CC–D		.84		.84
ZUNGD		.89		.89
MAACL–A		.57		.69
CC–A		.62		.75

Note. All models included estimation of correlation errors between related anxiety and depression subscales. The standardized parameter estimates are presented. All parameters were significant at the .05 level. ϕ = the correlation between latent constructs; λ = the loading of an observed indicator on a latent variable; BDI = Beck Depression Inventory; MAACL–D = Multiple Affect Adjective Check List, Depression subscale; MAACL–A = Multiple Affect Adjective Check List, Anxiety subscale; CC–D = Costello–Comrey Depression scale; CC–A = Costello–Comrey Anxiety scale; A-Trait = State–Trait Anxiety Inventory, trait version; A-State = State–Trait Inventory, state version; SCL–D = Symptom Checklist—90–Depression subscale; SCL–A = Symptom Checklist—90–Anxiety subscale; ZUNGD = Zung Depression Scale; ZUNGA = Zung Anxiety scale; DACL = Lubin Depression Adjective Checklist.

Table 2
Comparison of Fit Between One- and Two-Factor Models

Study/model	*df*	χ^2	*p*	CFI	MCFI
Dobson (1985)					
One-factor model	7	24.80	.001	.94	.38
Two-factor model	6	22.70	.001	.94	.42
Difference	1	2.10	*ns*		
Gotlib (1984)					
One-factor model	7	131.60	.001	.93	.69
Two-factor model	6	85.52	.001	.96	.80
Difference	1	46.08	.001		
Tanaka-Matsumi and Kameoka (1986)					
One-factor model	7	49.25	.001	.97	.74
Two-factor model	6	49.25	.001	.97	.74
Difference	1	0.00	*ns*		
Mendels, Weinstein, and Cochrane (1972)					
One-factor model	7	21.63	.001	.96	.80
Two-factor model	6	7.14	3.08	.997	.98
Difference	1	14.99	.001		

Note. CFI = comparative fit index; MCFI = modified comparative fit index.

the data well. Comparative normed-fit indices for all of the analyses are presented in Table 2; all indices were above .90 for analyses. Respective one- and two-factor models were comparable in their ability to account for the observed data from all data sets.

When competing models obtain similar normed-fit indices, the difference in fit may be too small to be detected if a null model is used in the calculation of the index. Some researchers have suggested that for the detection of more subtle differences, the CFI should be calculated using a baseline model that is less restricted than the null model (Mulaik et al., 1989; Sobel & Bohrnstedt, 1985). Accordingly, the hypothesized one- and two-factor models were compared with a baseline model in which two latent constructs were hypothesized and the loadings (λ) within each construct were constrained to be equal to one another (see Mulaik et al., 1989). The modified CFI coefficients (MCFI) presented in Table 2 indicate clearly that the two-factor models did not perform better than did the one-factor model in accounting for the variance in the Dobson (1985a) and Tanaka-Matsumi and Kameoka (1986) data sets. Like the difference chi-square statistic, however, the MCFI suggested that the two-factor model may have fit the data better in the Gotlib (1984) and Mendels et al. (1972) data sets.

One characteristic of normed-fit indices is that the fit of a model increases simply by freeing up parameters to be estimated (James, Mulaik, & Brett, 1982). Each additional parameter that is freed to be estimated in the analysis removes one constraint on the final solution, with the result that the reproduced data matrix will better fit the sample data matrix (Mulaik et al., 1989). Thus, a two-factor model may fit the data better than the one-factor model simply because one additional parameter is being estimated. This may account for the better fit of the two-factor model when compared with the one-factor model for the Mendels et al. (1972) and the Gotlib (1984) data sets.

Incremental normed-fit index (INFI). The INFI (Bentler, 1990) allowed a direct comparison of the one- and two-factor models. This index represents the relative improvement of the two-factor model over the one-factor model. If the expanded model provides a small improvement, the fit index will be close to 0; if the model represents a large improvement, the fit index will be closer to 1. The INFIs for the four sets of confirmatory factor analyses indicated that the two-factor models did not fit the data matrices better than did the one-factor models (INFI = .01, Dobson, 1985a; INFI = .02, Gotlib, 1984; INFI = .00, Tanaka-Matsumi & Kameoka, 1986; INFI = .04, Mendels et al., 1972). The INFI provided evidence that there was little improvement of the two-factor model over the one-factor model for the Mendels et al. (1972) and the Gotlib (1984) data sets.

Correlation Between Anxiety and Depression Latent Constructs

Further evidence that the anxiety and depression scales do not measure discriminant constructs comes from the size of the correlations between the latent variables of anxiety and depression in the two-factor confirmatory factor analyses. The correlation between latent variables should represent the true correlation between anxiety and depression constructs when corrected for attenuation due to the unreliability in the measured variables (Breckler, 1990). Analyses of all four data sets yielded large correlations between the latent constructs of anxiety and depression (ϕ = .96, Dobson, 1985a; ϕ = .92, Gotlib, 1984; ϕ = .999, Tanaka-Matsumi & Kameoka, 1986; ϕ = .82, Mendels et al., 1972). There is essentially no unique variance in the latent anxiety and depression constructs as estimated from the nonclinical data sets. Furthermore, although the correlation between latent anxiety and depression constructs is somewhat lower in the clinical sample than in the nonclinical samples, the correlation between constructs is high enough that one would have to carefully consider treating self-reported anxiety and depression as distinct in this data set.

Discussion

The evidence strongly indicated that there is little usefulness in discriminating between self-report measures of anxiety and depression in the nonclinical data sets. The two-factor models did not capture the structure of the self-report measures any better than did the one-factor models. Furthermore, correlations between the anxiety and depression latent constructs were so large that, even if one were to accept the two-factor model, there was little unique variance associated with the anxiety or the depression latent variable. Thus, these analyses suggested that the self-report measures of anxiety and depression used here do not measure unique mood constructs in nonclinical samples and should be considered measures of general distress rather than measures of anxiety and depression per se.

The confirmatory factor analyses of the Mendels et al. (1972) data provided some evidence that anxiety and depression self-report measures lack discrimination even when used with clinical samples. Although there was some disagreement among the fit indices, both the fit index involving a direct comparison of the models with reference to a null model and the correlation between latent constructs suggested that there was little usefulness in the two-factor model. Overall, most aspects of the confirmatory factor analyses of the Mendels et al. (1972) data set indicated that there is little discrimination in self-reported anxiety and depression in clinical samples. One reason for the mixed nature of the results from the Mendels et al. data set may be that this sample appears to have been heterogeneous. The publication did not give any information regarding the diagnostic status of the respondents included in the sample. The respondents may have had both depressive and anxiety disorders, in which case one would not expect to be able to distinguish these moods in this sample. The Mendels et al. sample may be representative of typical clinical patients, given the large degree of symptom comorbidity of anxiety and depression in clinical samples (for a review, see Katon & Roy-Byrne, 1991; Maser & Cloninger, 1990). It is important to note, however, that self-reported anxiety and depression have not been discriminable even in distinct diagnostic groups (e.g., Downing & Rickels, 1974; Lipman, 1982; Prusoff & Klerman, 1974).

Measurement of Anxiety and Depression

The results of the confirmatory factor analyses of the anxiety and depression ratings provided the strongest demonstration to date that the two moods cannot be meaningfully separated by self-report measures. There are two possible explanations for these results. Anxiety and depression may actually be distinct mood experiences but existing measures are incapable of tapping their distinct features. Alternatively, anxiety and depression may not be distinct mood experiences at all but merely two examples of general negative mood. These two possibilities will be presented in turn.

Anxiety and depression are distinct mood experiences, but existing measures are incapable of detecting the distinction. If one assumes that anxiety and depression are discriminant mood experiences, one might attempt to alter the measures so that they more validly assess distinct latent constructs. One possible method of increasing the discrimination in anxiety and depression self-report measures was suggested by L. A. Clark and Watson (1991). These authors have suggested that, in addition to sharing a nonspecific negative mood component, anxiety and depression can be distinguished because anxiety is associated with physiological hyperarousal, whereas depression is associated with low Positive Affect (PA), or the loss of pleasurable engagement with the environment (see also Tellegen, 1985; Watson, Clark, & Carey, 1988; Watson & Kendall, 1989). The high correlations between anxiety and depression measures may be due to the fact that most self-report measures of both primarily tap Negative Affect (NA). Differential measurement of the two moods may be improved if depression scales are revised to reflect the absence of high PA, whereas anxiety scales should be adjusted to reflect psychological arousal (L. A. Clark & Watson, 1991).

Although adjusting self-report measures of anxiety and depression in such a manner may seem sound, there is some uncertainty as to whether such adjustments will increase the discrimination of the two moods in self-report. First, there is preliminary evidence that ratings of anxiety and depression terms are in large part ratings of positivity or negativity and tend not to reflect the differences in the level of physiological arousal that differentiate anxiety from depression (Feldman, 1991). The perception of arousal seems to have a relatively minor impact on mood ratings when compared with valence (Feldman, 1991). It is not even clear that clinically anxious and depressed patients make use of autonomic arousal in reporting their mood (e.g., Leff, 1978).

Second, adding high-PA items to depression measures may not serve to discriminate anxiety and depression in self-report. High-PA words (e.g., *inspired* or *active*) are essentially positively valenced words that denote high arousal. Low ratings on these words represent negative valence and low arousal. Given that existing depression scales already include negatively valenced mood words, adding high-PA items amounts to adding words that will tap the absence of high arousal states (which may or may not be the same as low arousal states; Feldman, 1993). If mood ratings do not reflect differences in arousal states, then adding such words should not alter the correlations between anxiety and the revised depression scales. There is preliminary evidence to suggest that the correlations between several anxiety and depression self-report measures are equivalent regardless of whether the depression scales include high-PA items (Feldman, 1993). It is not clear whether revising self-report measures of anxiety to reflect physiological hyperarousal and measures of depression to reflect the absence of high PA will make the mood states more distinguishable in self-report.

It may be argued that merely adding PA items to depression scales is an oversimplification of the Clark and Watson (1991) model. It is the case that if depression scales are modified to primarily reflect PA, the correlations with anxiety measures decrease (Feldman, 1993). This is not surprising, given previous evidence that anxiety is unrelated to PA (Tellegen, 1985; Watson, Clark, & Tellegen, 1988). But such a modification implies that the construct of depressed mood is being redefined as the absence of PA. The validity of such a construct shift should be investigated before it is widely accepted (Feldman, 1993).

Another solution to the discrimination problem may involve the use of measures involving something other than self-report. Two obvious candidates are behavioral and physiological measures. Although these may be avenues worth exploring, the little research that has been done has yielded confusing results. Most research attempting to discriminate between anxiety and depression through behavioral measures has met with inconsistent success. For example, Higgins, Bond, Klein, and Strauman (1986) found that changes in writing speed were associated with experimentally induced increases in depression, but they found no change in writing speed for increases in anxiety. In contrast, findings from other studies (Strauman, 1988; Strauman & Higgins, 1987) indicated that a manipulation designed to increase anxiety was associated with changes in the total length of time spent verbalizing during a sentence completion task but that a manipulation designed to increase depression was not. The speed of verbal responses given to a sentence completion task was moderately related to manipulations designed to increase anxiety or depression in the Strauman and Higgins studies, but these findings were not replicated in Strauman's study. All results occurred alongside the finding that subjects did report increases in self-reported anxiety and depression in the appropriate experimental conditions.

Like behavioral measures, physiological measures have had little consistent success in discriminating between anxiety and depression. Strauman (Strauman, 1988; Strauman & Higgins, 1987) obtained evidence that skin conductance changes with the experience of anxiety and depression. Depression was related to decreases in arousal, whereas anxiety was related to increases in arousal. Unfortunately, Ekman, Levenson, and Friesen (1983) were unable to detect differences in skin conductance between anxiety and depression in respondents who were asked to construct facial prototypes and relive past emotional experiences.

In sum, self-report behavioral and physiological measures have not consistently discriminated between anxiety and depression. Possible discrimination may be increased by altering self-report scales in some fashion, but the efficacy of this approach remains an empirical question. If anxiety and depression are distinct psychological entities, they appear to be empirically indistinguishable as far as current measurement methods allow.

Anxiety and depression may be distinct mood states, but they may also covary to such a degree that they are rarely found alone. In the face of such high covariation, however, one wonders at the usefulness of treating anxiety and depression as separate constructs. Indeed, evidence from recent clinical studies calls into question whether depression and anxiety are discriminable as clinical entities. The evidence from treatment overlap studies, epidemiological studies, genetic studies, self-reported mood studies, and comorbidity studies, both at the level of clinical symptoms and at the level of syndrome diagnosis, all suggest the possibility that anxiety and depression may be caused by some unitary underlying vulnerability (for a review, see Maser & Cloninger, 1990). Indeed, some researchers are now even questioning the clinical validity of discriminating the two constructs. In this context, it is noteworthy that the pattern of findings described for the self-report ratings has also been obtained in clinicians' ratings of patients (see D. A. Clark et al., 1990; L. A. Clark, 1989; Stavrakaki & Vargo, 1986). High correlations and single-factor solutions typically characterize the research using clinically based ratings of anxious and depressed patients. A differentiation has been found in correlational studies (e.g., Beck et al., 1988; Riskind, Beck, Brown, & Steer, 1987) only when clinicians' ratings are made in the context of the diagnostic process, with an inclination toward differential diagnosis between anxiety and depression (L. A. Clark, 1989). The clinical evidence, then, also casts some doubt as to whether anxiety and depression are distinct psychological constructs.

Anxiety and depression are not distinct psychological states but merely exemplars of negative mood. If psychologists assume that anxiety and depression are simply two examples of general negative mood, then mood assessment should involve the measurement of more global mood states. Indeed, some theorists (e.g., Eysenck, 1970; Watson & Clark, 1984) have proposed personality constructs that treat manifestations of anxiety and depression as part of broader constructs as well as separate entities.

Researchers could begin by measuring the dimensions that characterize mood, rather than by measuring individual mood states. For example, researchers could make use of Russell's Affect Grid to assess valence (positivity or negativity) and the level of arousal (Russell, Weiss, & Mendelsohn, 1989), or they could opt for the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988), which provides measures of NA and PA. The advantage of the dimensional approach is that there is no assumption of the existence of anxiety and depression constructs that cannot be measured independently.

There is some uncertainty when choosing a dimensional strategy to measure mood. Researchers continue to debate about whether NA–PA or arousal–valence are the primary dimensions of mood. Because the models are rotational variants of each other, dimensionality may be chosen by theoretical considerations that need to be explicitly specified by the researcher. Interestingly, both models offer constructs that are associated with diagnostic criteria of anxiety or depression (Feldman, 1991). Preliminary evidence suggests that neither model is able to effectively differentiate anxiety and depression in self-report, however (Feldman, 1991). Thus, the ability to distinguish between anxiety and depression cannot be cited as a reason to prefer one rotational variant to the other.

Future Directions

Although both laypeople and psychologists strongly believe that anxiety and depression are distinct entities, the results of these studies suggest that they are not separable with self-report measures. Psychologists, however, have tended to proceed with their research as if the distinction were still likely to be correct. Although it can be said that the present analyses did not include every conceivable self-report measure available, the scales used in these analyses are representative of those that researchers tend to use when measuring anxiety and depression.

The discrimination problem is not specific to measures of anxiety and depression alone. There is little evidence to suggest that self-report measures of other types of negative mood (e.g., anger) validly assess constructs that are distinct from anxiety and depression (e.g., Polivy, 1981; Zuckerman et al., 1983).

In the broadest context, the implication of this investigation is clear. Without discriminant measurement of anxiety and depression, the results of previous empirical investigations are difficult to interpret. One cannot assume a relationship between cognitive, behavioral, biological, and personality variables of interest on the one hand and anxiety or depression on the other, when the reports of both negative moods may be elevated. Furthermore, any theory based on this discrimination may be unverifiable to the extent that studies rely on subjective reports of anxiety and depression. Much of the previous research conducted to date using self-reports of anxiety and depression becomes difficult to interpret given that anxiety and depression self-report scales merely measure general negative mood.

References

Anderson, J. C., & Gerbing, A. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49,* 155–173.

Beck, A. T. (1976). *Cognitive theory and the emotional disorders.* Madison, CT: International Universities Press.

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

Beck, A. T., Ward, C. M., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Bentler, P. M. (1990). Comparative fit indexes in structural modeling. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M., & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin, 105,* 467–477.

Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin, 107,* 260–273.

Bush, L. E. (1973). Individual differences multidimensional scaling of adjectives denoting feelings. *Journal of Personality and Social Psychology, 25,* 50–57.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Clark, D. A., Beck, A. T., & Stewart, B. (1990). Cognitive specificity and positive–negative affectivity: Complementary or contradictory views on anxiety and depression? *Journal of Abnormal Psychology, 99,* 148–155.

Clark, L. A. (1989). The anxiety and depressive disorders: Descriptive psychopathology and differential diagnosis. In P. C. Kendall & D. Watson (Eds.), *Anxiety and depression: Distinctive and overlapping features* (pp. 83–129). San Diego, CA: Academic Press.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316–336.

Costello, C. G., & Comrey, A. L. (1967). Scales for measuring depression and anxiety. *Journal of Psychology, 66,* 303–313.

Cudeck, R. (1989). Analyses of correlation matrices using covariance structure models. *Psychological Bulletin, 105,* 317–327.

Davitz, J. R. (1969). *The language of emotion.* San Diego, CA: Academic Press.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). The SCL-90: An outpatient psychiatric rating scale. *Psychopharmacology Bulletin, 9,* 13–28.

Dobson, K. S. (1985a). An analysis of anxiety and depression scales. *Journal of Personality Assessment, 49,* 522–527.

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

Downing, R. W., & Rickels, K. (1974). Mixed anxiety–depression: Fact or myth? *Archives of General Psychiatry, 30,* 312–317.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science, 221,* 1208–1210.

Eysenck, H. (1970). *The structure of human personality.* London: Methuen.

Feldman, L. A. (1991). *On the failure to differentiate anxiety and depression in self-report.* Manuscript submitted for publication.

Feldman, L. A. (1993). *The role of positive affect in discriminating self-reported depression from anxiety.* Unpublished manuscript.

Gerbing, D. W., & Anderson, J. C. (1985). The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis. *Multivariate Behavioral Research, 20,* 255–271.

Gotlib, I. H. (1984). Depression and general psychopathology in university students. *Journal of Abnormal Psychology, 93,* 19–30.

Gotlib, I. H., & Meyer, J. P. (1986). Factor analysis of the Multiple Affect Adjective Check List: A separation of positive and negative affect. *Journal of Personality and Social Psychology, 50,* 1161–1165.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances.* Baltimore: Johns Hopkins University Press.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319–340.

Higgins, E. T., Bond, R. N., Klein, R., & Strauman, T. (1986). Self-discrepancies and emotional vulnerability: How magnitude, accessibility, and type of discrepancy influence affect. *Journal of Personality and Social Psychology, 51,* 5–15.

Hull, J. G. (1989). *LISREL primer: An introduction to structural modeling with latent variables.* Unpublished manuscript.

Izard, C. E. (1977). *Human emotions.* New York: Plenum Press.

James, L. R., Mulaik, S. A., & Brett, J. (1982). *Causal analysis: Models, assumptions and data.* Beverly Hills, CA: Sage.

Jöreskog, K. G., & Sörbom, D. G. (1984). *LISREL VI.* Mooresville, IN: Scientific Software.

Katon, W., & Roy-Byrne, P. P. (1991). Mixed anxiety and depression. *Journal of Abnormal Psychology, 100,* 337–345.

Klerman, G. L. (1977). Anxiety and depression. In G. D. Burrows (Ed.), *Handbook of studies on depression* (pp. 49–68). New York: Excerpta Medica.

Leff, J. P. (1978). Psychiatrists' versus patients' concepts of unpleasant emotions. *British Journal of Psychiatry, 133,* 306–313.

Lipman, R. S. (1982). Differentiating anxiety and depression in anxiety disorders: Use of ratings scales. *Psychopharmacology Bulletin, 18,* 69–77.

Lubin, B. (1967). *Depression Adjective Checklists: Manual.* San Diego, CA: Educational and Industrial Testing Service.

Maser, J. D., & Cloninger, C. R. (1990). *Comorbidity of mood and anxiety disorders.* Washington, DC: American Psychiatric Press.

Mendels, J., Weinstein, N., & Cochrane, C. (1972). The relationship between depression and anxiety. *Archives of General Psychiatry, 27,* 649–653.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105,* 430–445.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge, England: Cambridge University Press.

Pennebaker, J. W. (1982). *The psychology of physical symptoms.* New York: Springer-Verlag.

Polivy, J. (1981). On the induction of emotion in the laboratory: Discrete moods or multiple affect states? *Journal of Personality and Social Psychology, 41,* 803–817.

Prusoff, B., & Klerman, G. L. (1974). Differentiating depressed from anxious neurotic outpatients. *Archives of General Psychiatry, 30,* 302–309.

Riskind, J. H., 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.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57,* 493–502.

Sobel, M. E., & Bohrnstedt, G. W. (1985). Use of null models in evaluating the fit of covariance structure models. In N. B. Tuma (Ed.), *Sociological methodology* (pp. 152–178). San Francisco: Jossey-Bass.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *STAI—Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Stavrakaki, C., & Vargo, B. (1986). The relationship of anxiety and depression: A review of the literature. *British Journal of Psychiatry, 149,* 7–16.

Strauman, T. J. (1988). Self-discrepancies in clinical depression and social phobia: Cognitive structures that underlie emotional disorders? *Journal of Abnormal Psychology, 98,* 14–22.

Strauman, T. J., & Higgins, E. T. (1987). Automatic activation of self-discrepancies and emotional syndromes: When cognitive structures influence affect. *Journal of Personality and Social Psychology, 53,* 1004–1014.

Tanaka, J. S. (1987). "How big is big enough"? Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58,* 134–146.

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.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maswer (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.

Wallbott, H. G., & Scherer, K. R. (1989). Assessing emotion by questionnaire. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (Vol. 4, pp. 55–82). San Diego, CA: Academic Press.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465–490.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97,* 346–353.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Watson, D., & Kendall, P. C. (1989). Understanding anxiety and depression: Their relation to negative and positive affect states. In P. C. Kendall & D. Watson (Eds.), *Anxiety and depression: Distinctive and overlapping features* (pp. 3–26). San Diego, CA: Academic Press.

Zuckerman, M., & Lubin, B. (1965). *Manual for the Multiple Affect Adjective Check List.* San Diego, CA: Educational and Industrial Testing Service.

Zuckerman, M., Lubin, B., & Rinck, C. M. (1983). Construction of new scales for the Multiple Affect Adjective Check List. *Journal of Behavioral Assessment, 5,* 119–129.

Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry, 12,* 63–70.

Zung, W. W. K. (1971). A rating instrument for anxiety disorders. *Psychosomatics, 12,* 371–379.

Received November 22, 1991
Revision received June 8, 1992
Accepted July 7, 1992 ■