Psychological Assessment
2002, Vol. 14, No. 2, 135–145

1040-3590/02/$5.00 DOI: 10.1037//1040-3590.14.2.135

The Latent Structure of Analogue Depression: Should the Beck Depression Inventory Be Used to Classify Groups?

Ayelet Meron Ruscio
Pennsylvania State University

John Ruscio
Elizabethtown College

Research on depression is often conducted with analogue samples that have been divided into depressed and nondepressed groups using a cutoff score on the Beck Depression Inventory (BDI). Although the relative merits of different cut scores are frequently debated, no study has yet determined whether the use of *any* cut score is valid, that is, whether the latent structure of BDI depression is categorical or dimensional in analogue samples. The BDI responses of 2,260 college students were submitted to 3 taxometric procedures whose results were compared with those of simulated data sets with equivalent parameters. Analyses provided converging evidence for the dimensionality of analogue depression, arguing against the use of the BDI to classify analogue participants into groups. Analyses also illustrated the notable impact of pronounced skew on taxometric results and the value of using simulated comparison data as an interpretive aid.

A vast number of investigations into the characteristics and correlates of depression—perhaps even the majority of such investigations—have been conducted with analogue, or college student, samples (Tennen, Hall, & Affleck, 1995; Vredenburg, Flett, & Krames, 1993). These studies have tended to assess depressive symptoms through self-report questionnaires, the most popular being the Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979). Because the BDI provides a continuous severity score rather than a dichotomous diagnostic decision, researchers typically identify participants as depressed or nondepressed on the basis of a given cutoff point along the score distribution, then test for differences across the resultant groups (Vredenburg et al., 1993). Some investigators even divide the BDI score distribution into three or more groups, each defined by a different level of depressive severity (e.g., Andersen, Spielman, & Bargh, 1992; Pelham, 1991).

Vredenburg et al. (1993) have observed that there have been few empirical attempts to identify appropriate cut score criteria for the BDI, leaving investigators to choose cutoff values largely on the basis of speculation or convention when forming groups for depression research. Indeed, although the modal value used for group classification is a BDI score of 10, scores between 11 and 16 are used by some researchers to increase the generalizability of their findings to clinical samples (see Tennen, Eberhardt, & Affleck, 1999; Tennen et al., 1995), and scores as low as 9 (e.g., Pinkley, Laprelle, Pyszczynski, & Greenberg, 1988) and as high as 24 (e.g., Marsh & Weary, 1994) are occasionally used. There have been numerous criticisms of the BDI cut scores typically used in analogue research, almost all concerning the degree to which the "depressed" group formed by such cuts corresponds to individuals diagnosed with major depressive disorder in clinical settings. Critics have noted that the BDI is a measure of syndromal depression that was not designed to yield diagnoses (Deardorff & Funabiki, 1985); that a cut score as low as 10 may classify as "depressed" individuals who experience very few symptoms of depression and who do not endorse its central features of depressed mood and loss of interest (Ingram & Hamilton, 1999; Tennen et al., 1995); that low to moderate cut scores are relatively nonspecific to depression and may instead identify those with anxiety disorders, substance disorders, or general distress (Coyne, 1994; Deardorff & Funabiki, 1985; Kendall & Flannery-Schroeder, 1995); and that the cutoffs used to designate individuals as nondepressed are often inappropriate (Kendall & Flannery-Schroeder, 1995; Kendall, Hollon, Beck, Hammen, & Ingram, 1987). Moreover, whereas some critics lament the obstruction of cross-study comparisons caused by the inconsistent selection of cut scores (e.g., Kendall et al., 1987), others protest researchers' selection of such scores on the basis of popularity and convenience, without consideration for such fundamental factors as prior probabilities, the sensitivity and specificity of the score within the population of interest, and the nature of the question under study (Tennen et al., 1999).

In light of these and other criticisms, Vredenburg et al. (1993) have called for a systematic evaluation of cut scores, noting that "the importance of this research is underscored by the possibility that some researchers in this area will continue to use a measure such as the BDI as the primary or sole means of group assignment" (p. 339). Indeed, continued widespread use of the BDI to classify depression suggests that empirical attention to the issue of cut scores is critical. However, before research can address the ques-

Ayelet Meron Ruscio, Department of Psychology, Pennsylvania State University; John Ruscio, Department of Psychology, Elizabethtown College.

Portions of this article were presented at the annual meeting of the American Psychological Society, Toronto, Ontario, Canada, June 2001. This research was supported by National Research Service Award MH12675 to Ayelet Meron Ruscio from the National Institute of Mental Health. We thank T. D. Borkovec and Louis G. Castonguay for their valuable comments on drafts of this article.

Correspondence concerning this article should be addressed to Ayelet Meron Ruscio, Department of Psychology, Pennsylvania State University, 429 Bruce V. Moore Building, University Park, Pennsylvania 16802-3104. E-mail: amr205@psu.edu

tion of which cutoff value is best, it must first determine whether *any* cut along the BDI distribution validly distinguishes nonarbitrary groups. Although the correspondence of BDI-diagnosed analogue samples to depressed clinical samples is clearly important, it may be argued that the appropriateness of dichotomizing the BDI score distribution largely depends on whether underlying groups exist, irrespective of their similarity to groups found in clinical settings (J. Ruscio & Ruscio, 2002). The appropriateness of making a cut is not a matter of analytic convenience or of the relationship between analogue and clinical depression; rather, it is a question of the fundamental structure of analogue depression as measured by the BDI. It asks whether there is, in nature, a qualitative boundary separating depressed from nondepressed individuals in analogue samples, or whether analogue depression lies along a latent continuum uninterrupted by qualitative breaks.[1]

The question of latent structure is one that concerns not only the BDI but the broader field of depression, in which debates persist over the nature of the boundary separating major depression from milder depressive states (Compas, Ey, & Grant, 1993; Grove & Andreasen, 1989). Although major depression is currently assessed and diagnosed within a categorical framework that at least implies a qualitative difference between disorder and normality (*Diagnostic and Statistical Manual of Mental Disorders, 4th ed.; DSM–IV*; American Psychiatric Association, 1994), there is growing evidence that depression may be better represented by a latent dimension that differs only quantitatively from milder mood states (Flett, Vredenburg, & Krames, 1997). Studies have found that individuals whose depressive symptoms do not qualify for a diagnosis of major depressive disorder nevertheless experience significant functional impairment (Broadhead, Blazer, George, & Tse, 1990; Johnson, Weissman, & Klerman, 1992), have elevated rates of major depression among their relatives (Kendler & Gardner, 1998; Sherbourne et al., 1994), and are more likely to experience a subsequent major depressive episode (Horwath, Johnson, Klerman, & Weissman, 1992; Kendler & Gardner, 1998), raising questions about the clinical importance of subthreshold depressive states and the discreteness of their boundary with major depressive disorder. These questions are further fueled by indications that increases in the number, severity, and duration of depression symptoms are associated with a corresponding linear increase in impairment, comorbidity, and familial depressive episodes, with no apparent discontinuity at the *DSM* diagnostic boundary (Judd, Akiskal, & Paulus, 1997; Kendler & Gardner, 1998; Kessler, Zhao, Blazer, & Swartz, 1997).

These findings suggest that subthreshold depressive symptoms and diagnosed depression lie along a single continuum of depressive severity and hint that the structure of measures that assess depression—such as the BDI—may be similarly continuous in nature. However, although these studies provided strong evidence for the continuity of depressive symptoms at the manifest level, they did not directly examine the structure of syndromal depression at the latent level. To this end, we conducted a study using Meehl's (1995) taxometric method—an approach specifically designed to distinguish latent groups (or taxa) from continua—to examine the latent structure of major depression (J. Ruscio & Ruscio, 2000). Taxometric analyses were performed in two large clinical samples with high rates of severe depression, using variables capturing the major cognitive and somatic features of the disorder. Results converged on a dimensional solution, suggesting that major depression is continuous with milder depressive states at the latent level.

Although this research was primarily concerned with major depression, its findings suggest that depression may be similarly continuous at lower levels of symptom severity, such as those typically found in analogue samples. Furthermore, because the BDI was one of the measures that revealed the dimensional structure of depression in this clinical study, it seems particularly important to extend this research to the analogue samples in which the BDI is so often used to classify individuals into groups. If there are meaningful groups underlying the BDI score distribution, such research will help to identify the BDI cut score that best distinguishes these groups within an analogue population (Meehl & Rosen, 1955). If, on the other hand, the BDI corresponds to a latent dimension, the primary question will shift away from *where* to cut the score distribution to *whether* a cut should be made, and why. Although researchers may use cut scores for purely pragmatic reasons rather than because they believe that their chosen score meaningfully separates natural groups, the use of cut scores in the absence of such groups could significantly impede our understanding of analogue depression. This is because if BDI depression is dimensional in analogue samples, any division of these samples into depressed and nondepressed components would necessarily construct an arbitrary boundary that not only would be theoretically misleading, but might also lead to a considerable loss of statistical power and a resulting obfuscation of important effects. These concerns are by no means specific to the BDI or to analogue populations but reflect a lack of consensus within the broader field of psychology about the circumstances under which the categorization of assessment instruments is appropriate. Although Cohen (1983) observed almost 2 decades ago that the spurious dichotomization of a continuous score distribution is commensurate to throwing away 36% or more of a sample, such dichotomization remains a common practice in psychological research. However, if we wish to rapidly accumulate knowledge about the critical construct of depression, it is important not only to determine the latent structure of this construct in different populations but to use such structural knowledge to enhance the power of its assessment and investigation.

In the present study, we examined the latent structure of analogue depression, as measured by the BDI, in a large college sample comparable to those typically used in depression research. We assessed structure through multiple analytic procedures within the taxometric method. Because researchers most often use a BDI cut score of 10 to classify depression in analogue samples and because a cut score of 16 has been specifically recommended to distinguish individuals with syndromal depression from those with milder dysphoria (Kendall et al., 1987; Vredenburg et al., 1993), we designed the analyses so as to be particularly sensitive to qualitative boundaries at these values.

[1] To underscore the fact that endorsement of BDI symptoms by analogue participants does not necessarily reflect a diagnosis of major depression, we refer to such symptom endorsement as *analogue depression* or *BDI depression* throughout the article.

Method

Participants and Measure

Participants were 2,260 undergraduate students (69% women and 31% men) enrolled in introductory psychology at a large northeastern university. These students completed the BDI as part of a larger questionnaire battery in exchange for extra credit. The BDI is the most widely used self-report measure of depression (Katz, Shaw, Vallis, & Kaiser, 1995), with excellent psychometric properties in both clinical and nonclinical samples (Beck & Steer, 1993; Beck, Steer, & Garbin, 1988; Lips & Ng, 1985). Each of its 21 items assesses a specific symptom of depression, asking respondents to rate, on a scale of 0 to 3, the intensity with which they have experienced that symptom during the past week. In the present sample, BDI total scores ranged from 0 to 41 and were positively skewed (mode = 0, *Mdn* = 5, $M = 6.37$). The mean, standard deviation, and internal consistency of BDI scores in the present sample ($M = 6.37$, $SD = 6.26$, $\alpha = .87$) were comparable to those reported in a recent psychometric evaluation of the BDI in a large undergraduate sample ($M = 7.42$, $SD = 6.67$, $\alpha = .89$; Dozois, Dobson, & Ahnberg, 1998), suggesting that our sample was similar to analogue samples that typically complete the BDI.

Taxometric Analyses

Meehl and his colleagues (e.g., Meehl, 1995, 1999; Meehl & Yonce, 1994, 1996; Waller & Meehl, 1998) have pioneered a method for determining the latent structure of a psychological construct. The various procedures within this taxometric method search for orderly statistical relations between indicators of the construct that are indicative of a qualitative boundary between two latent groups, traditionally referred to as the *taxon* (e.g., depressed) and *complement* (e.g., nondepressed). Provided that the available data are sufficiently sensitive and specific to the putative boundary under investigation, the taxometric method is neutral in its evaluation of this boundary. In other words, rather than beginning with a default position or null hypothesis concerning latent structure, results are examined for evidence consistent with either taxonic or dimensional structure. Moreover, rather than relying on traditional significance tests, conclusions are based on the convergence of results across multiple, quasi-independent analytic procedures. Each procedure serves as a consistency check for the results of the others, with confidence in a structural solution increasing as each additional test is passed. Because the taxometric method hinges critically on converging evidence, we used three mathematically distinct taxometric procedures to examine the latent structure of analogue depression.

Although Monte Carlo studies had indicated that taxometric procedures work well in the presence of indicator skew (Cleland & Haslam, 1996; Haslam & Cleland, 1996), these studies simulated a relatively small amount of skew. Because our BDI data, as expected, were strongly positively skewed, we took several steps to facilitate accurate interpretation of our taxometric results. First, on the basis of existing research and the logic of the taxometric procedures, we specified a priori the anticipated effects of pronounced skew on the results of each procedure. These hypothesized effects and their deductive basis are summarized in our description of each procedure below. Second, to provide a preliminary test of the impact of pronounced skew on taxometric results, as well as to ensure that our data were appropriate for taxometric analysis, we simulated taxonic and dimensional comparison data that matched the parameters of our sample. Analyses performed in these simulated data sets were then compared with those of our BDI data to aid interpretation of results.

MAMBAC. The first procedure that we used was MAMBAC (mean above minus below a cut; Meehl & Yonce, 1994), which is based on the premise that if two latent taxa exist, there must be an optimal cutting score for distinguishing between them. That is, if an indicator validly separates two latent taxa, there must be a particular cutting score on this indicator that will minimize the number of false-positive and false-negative classifications of cases into the taxa. In the absence of latent taxa, such an optimal cutting score will not exist.

MAMBAC requires two valid indicators that are correlated in the total sample. One of these indicators is treated as the *input* and is placed along the *x*-axis of the MAMBAC graph. Cases are sorted by their scores (from lowest to highest) on this input indicator. Then, the mean score on the other (*output*) indicator for all cases falling below the input cut is subtracted from the mean score of all cases falling above the input cut. This subtraction is repeated for all possible cutting scores along the *x*-axis (i.e., between each successive case in the data set), and each mean difference is plotted as the corresponding *y*-value on the MAMBAC graph. Thus, MAMBAC involves plotting mean differences on the output indicator above, minus below, cutting scores along the input indicator. The shape of the resulting curve indicates whether the latent structure is taxonic or dimensional. Taxonic structure yields a peaked MAMBAC curve: As the sliding cutting score approaches the optimal value for differentiating the latent groups, the mean difference between cases below and above the cut steadily increases toward a peak, then declines as the optimal value is passed. In contrast, the prototypical MAMBAC curve for dimensional structure is not peaked, but concave, often curving upward at one or both ends.

Although taxonic and dimensional MAMBAC curves are usually readily distinguished, indicator skew can complicate matters by "tilting" the shapes of the curves. For example, consider a hypothetical analysis with positively skewed indicators ranging in value from 0 to 100 with a mean of 20. At the left end of the MAMBAC curve, the input cuts will yield a mean difference of approximately 20, because virtually all cases will lie above the cut ($M \approx 20$) while the few lying below the cut will quite likely score very low on the output indicator ($M \approx 0$) because of the correlation between indicators. At the right end of the MAMBAC curve, the input cuts will yield a mean difference of approximately 80, because virtually all cases will lie below the cut ($M \approx 20$), whereas the few cases lying above the cut will tend to score very highly on the output indicator ($M \approx 100$). Thus, regardless of the underlying structure, the *y*-values of this MAMBAC curve will tend to increase from about 20 at the left end of the graph to about 80 at the right. Relative to this tilted curve, taxonic data should still produce a discernible peak, whereas dimensional data should not.

MAMBAC provides structural information not only by its shape but also by its estimate of the taxon base rate in the sample (Meehl & Yonce, 1994). This estimate is based on the ratio between MAMBAC values at the two endpoints of the curve. For example, a taxon whose base rate is approximately .50 will yield a MAMBAC curve with a central peak and comparably low endpoints. In the presence of a smaller taxon, the peak will be deflected toward the right, yielding a right endpoint that is higher than the left. By contrast, Meehl and Yonce (1994) demonstrated that dimensional structure produces roughly symmetrical MAMBAC curves whose endpoints are quite similar, generating base-rate estimates hovering around .50. It is important to note, however, that Meehl and Yonce (1994) used indicators that were simulated with no skew. If indicator skew was introduced, we would expect the consequent tilting of the MAMBAC curves to have predictable effects on base-rate estimates, regardless of the structural solution. Specifically, positive skew would be expected to increase the right endpoint of the curve relative to the left, thus reducing the base-rate estimate to a level similar to that of a low base-rate taxon.

MAXEIG. The second procedure to be used was MAXEIG (maximum eigenvalue; Waller & Meehl, 1998), a powerful multivariate taxometric procedure that permits the simultaneous use of *all* available indicator variables in each analysis. Like the more widely used MAXCOV (maximum covariance; Meehl & Yonce, 1996) procedure, MAXEIG tests whether indicators of a given construct covary because of a mixture of two underlying groups (taxonic latent structure) or loadings on a latent factor (dimensional latent structure). To distinguish between these possibilities, the association between indicators is explored within ordered subsets of cases in the sample.

Where latent groups exist, relatively pure subsamples of either taxon members (scoring high on each of the indicators) or complement members

(scoring low on each of the indicators) should reveal little association between indicators. In contrast, indicator association should be high in subsamples that contain a mixture of groups, because of their inclusion of both systematically high- and low-scoring cases. The strength of this association should reach a maximum in the subsample containing an equal mixture of taxon and complement members. By contrast, dimensional latent structure should produce a degree of association between indicators that remains fairly constant across subsets of a sample. This is because, in the absence of latent taxa, there is no reason for the strength of association to systematically vary. Thus, by examining the degree of association between indicators across ordered subsets of a sample, the latent structure of the underlying construct is revealed.

In MAXEIG, one indicator is selected to be the input and placed along the *x*-axis of the graph; all of the remaining indicators are used as outputs. Subsamples of cases are formed by sorting the sample according to scores on the input indicator and dividing the input into a series of overlapping *windows*. For example, the first window might contain the 100 lowest scoring cases on the input indicator, the second window might remove the lowest scoring 10 cases and replace them with the next 10 cases at the upper end of the window, and so forth. In this example (as in all of our analyses), adjacent sliding windows overlap 90% with one another. Within each window, the association between indicators is represented by the eigenvalue of the first principal factor that is derived from the covariance matrix of output indicators. The higher the covariance between indicators in the window, the higher the eigenvalue will be.[2] By plotting the eigenvalue for each window, one obtains a MAXEIG curve whose shape suggests the latent structure of the data. Taxonic structure yields a curve that peaks near the window containing a roughly equal mixture of taxon and complement members. Dimensional structure yields a relatively flat curve.

Indicator skew may also influence the shape of MAXEIG curves. Because the number of cases within each window is held constant, skew would be expected to produce a differential restriction of range on output indicators (and thus a differential reduction in the association between indicators) across the sliding windows. When a sliding window falls near the tail of a skewed distribution, it will include cases with considerably different output scores; when a sliding window falls near the body of a skewed distribution, it will contain less variation in scores. Thus, eigenvalues should be reduced in windows falling at the body of a skewed distribution relative to windows falling at the tail. For positively skewed indicators, the result should be an increasing MAXEIG curve that is "tilted" up toward the right, regardless of latent structure.

Fortunately, there is an additional test that can help to distinguish the upward-sloping MAXEIG curves of a low base-rate taxon from those of a latent dimension identified by positively skewed indicators. The *inchworm consistency test* was introduced by Waller and Meehl (1998) to help resolve the ambiguity of MAXEIG curves that peak at the endpoint of the graph. The test works by systematically increasing the number of windows that are included in the MAXEIG analysis. If the right-end peak is due only to the positive skew of the indicators, the peak will remain despite an increase in the number of windows. If, irrespective of skew, the peak is caused by a taxon so small that taxon members are outnumbered by complement members even in the rightmost windows of the graph, the peak should level off and then taper down as the number of windows is increased. This is because as the number of windows increases, the sample size within each window decreases, enabling one or more windows toward the end of the curve to surpass the point at which taxon members constitute 50% of the subsample. Once this point is passed, eigenvalues begin to decline once again, forming a more defined taxonic peak.

Following a procedure described by Waller and Meehl (1998), each MAXEIG curve can be used to estimate the taxon base rate. The midpoint of the window that yields the maximum eigenvalue is the *hitmax* for the input indicator, the value that maximizes the overall hit rate of classification by best distinguishing putative taxon and complement members. The proportion of cases above the hitmax value provides an estimate of the taxon base rate in the sample. Thus, centrally peaked curves suggest a taxon with a base rate of about .50, and to the extent that a peak is deflected toward the right, the base-rate estimate will be lower. Dimensional latent structure will not yield a consistent peak; rather, as the maximum eigenvalue is determined largely by sampling error, base-rate estimates tend to fluctuate markedly from curve to curve. Because positively skewed indicators are expected to produce upward-sloping MAXEIG curves, their base-rate estimates should be deflated regardless of latent structure.

L-Mode. The third procedure that we used was L-Mode (latent mode; Waller & Meehl, 1998). Like MAXEIG, L-Mode is powerful because it permits the simultaneous use of all available indicator variables. L-Mode works by factor analyzing all available indicators, then examining the distribution of estimated true scores on the first principal factor yielded by this analysis. Whereas the manifest distributions of indicator variables can be obscured by measurement error, L-Mode's reliance on estimated true scores minimizes the potentially obfuscating effects of such error. Taxonic data yield a bimodal score distribution, whereas dimensional data yield a unimodal score distribution. As with MAMBAC and MAXEIG, indicator skew would be expected to alter the shape of an L-Mode graph, in that scores on the first principal factor of a factor analysis performed with skewed indicators will evidence a skewed distribution.

Several estimates of the taxon base rate can be calculated for a given L-Mode graph (Waller & Meehl, 1998). One estimate is derived from the location of each latent mode. If these two estimates agree well with one another, they can be averaged to provide a more reliable value. A third estimate is derived from an empirical classification of cases. On the basis of its profile of scores on the indicator variables, each case in the sample can be classified as a taxon or a complement member. The proportion of cases classified into the taxon is then taken as another estimate of the taxon base rate.

Estimating the Taxon Base Rate

We have described how the graphical output of each taxometric procedure can be used to estimate the base rate of taxon membership in the sample. In the present study, we estimated the taxon base rate for each curve and examined the agreement among these estimates as an added consistency check of the structural solution suggested by the curves themselves. If a latent taxon underlies BDI scores in this analogue population, independent estimates of its base rate in the present sample should agree with one another. Failure of these estimates to cohere around a consistent value would instead suggest that analogue depression is a dimensional construct.

Simulated Comparison Data

Although taxometric analyses can distinguish taxonic from dimensional latent structure in the presence of indicator skew (Cleland & Haslam, 1996; Haslam & Cleland, 1996), the shape of the resulting curves may be affected in the ways described above. We therefore constructed two simulated data sets, one taxonic and one dimensional, to facilitate interpretation of the taxometric curves.[3]

The simulated dimensional data set was created by generating 2,260 cases of 10 random, normally distributed variables with shared loadings on

[2] For readers more familiar with MAXCOV, this procedure is conceptually similar. Whereas MAXCOV computes covariances between two output indicators within nonoverlapping intervals along an input indicator, MAXEIG computes eigenvalues between multiple output indicators within overlapping windows along an input indicator.

[3] Because sampling error has a negligible influence on analyses with samples this large, we simulated only one data set for each latent structure. As will be seen shortly, this afforded an unambiguous interpretation of our obtained curves.

a common latent factor. Each variable was then skewed and scaled like our BDI items by replacing the lowest 1,706 values with a score of 0, the next lowest 452 values with a score of 1, the next lowest 71 values with a score of 2, and the remaining 31 values with a score of 3, to match the average distribution of the BDI items in our sample. Factor loadings were set so that the final mean intercorrelation of the 10 simulated indicators matched that of the actual BDI indicators selected for analysis ($r = .34$; see below).

The simulated taxonic data set also began with the generation of 2,260 cases of 10 random, normally distributed variables. However, in contrast to the shared factor underlying the simulated dimensional data set, taxa were created by adding a constant (one matching the estimated validity of the BDI indicators) to each variable for a subset of cases that matched the proportion of our sample falling at or above a BDI score of 10 (.23; see below). Thus, each variable was constructed to validly separate the taxon from the complement. All variables were then skewed and scaled precisely as in the dimensional data set. The degree of taxonic separation was set so that the final intercorrelation of the 10 simulated indicators once again matched that of the actual BDI indicators. Therefore, the only difference between these two simulated data sets was their underlying structure (one latent dimension vs. two latent taxa); their manifest distributions, intercorrelations, and all other parameters and operations were held constant.

Results

Taxometric analyses were first performed in the full sample and then replicated separately for men and for women. Because each series of analyses yielded a highly similar pattern of results, results will be presented for the sample as a whole.[4]

Item Selection

Following conventional practice, the sample was divided into two parts, separating cases with BDI total scores of 0 to 9 ($n = 1{,}748$) from cases with total scores of 10 or above ($n = 512$). Items were then selected on the basis of their ability to validly distinguish between the putative depressive types, estimated by the mean difference between the two constructed groups divided by the within-group standard deviation (the average of the two groups, weighted by their sizes). Ten of the 21 BDI items achieved separations in excess of Meehl's (1995) recommended threshold of 1.25σ. To ensure that nuisance covariance was tolerably low, we computed correlations between these items within each of the two groups. Among the cases with BDI total scores less than 10, correlations ranged from $-.04$ to .31 ($M = .08$); among the cases with BDI total scores greater than or equal to 10, correlations ranged from .04 to .37 ($M = .18$). Because these separation and nuisance covariance estimates fell well within the levels recommended by Meehl (1995), we used these 10 BDI items as indicators in subsequent taxometric analyses.

The selected items—3, 4, 6, 7, 8, 9, 10, 12, 13, and 14—covered a variety of depressive features, including anhedonia, frequent crying, loss of interest in others, feelings of worthlessness and self-disgust, excessive self-criticism, impaired decision making, feelings of being punished, distorted body image, and suicidal ideation. Although we had intended to select a second set of indicators that optimally separated cases at a BDI cut score of 16, 9 of the 10 highest validity items at this cut score were identical to those obtained for a cut score of 10 (Item 4 would have been replaced by Item 5). Because the resulting taxometric curves would have been indistinguishable from those of our first series, we report only one series of taxometric analyses.

MAMBAC Analyses

Because individual BDI items (scaled from 0 to 3) possess too little variation to reliably sort cases, input indicators were computed by summing multiple BDI items. MAMBAC was performed using each of the 10 previously selected BDI items as output indicators in turn, with the sum of the remaining 9 items serving as the input indicator (program code for MAMBAC appears in the Appendix). All 10 curves sloped upward sharply, with base-rate estimates ranging from .07 to .21 ($M = .15$, $SD = .05$). However, although MAMBAC curves were tilted to the right due to the positive skew of the indicators (see Figure 1, top row), they were virtually identical to those obtained with the simulated skewed dimensional data (middle row) and quite unlike the tilted yet peaked curves obtained with the simulated skewed taxonic data (bottom row). Moreover, the base-rate estimates of the simulated dimensional data ($M = .15$) closely matched those of the BDI data, whereas the base-rate estimates of the simulated taxonic data ($M = .20$) more closely approximated the simulated taxon base rate of .23.

MAXEIG Analyses

As with MAMBAC, item combination was required to generate input indicators with sufficient variation to reliably sort cases. MAXEIG was thus performed using the 10 selected BDI items as output indicators, with the sum of the 11 unselected BDI items serving as the input indicator (see Waller & Meehl, 1998, for MAXEIG program code). The MAXEIG curve sloped upward toward the right, again influenced by the positive skew of the indicators. To determine whether this curve reflected a latent dimension or a low base-rate taxon, we conducted the inchworm consistency test by running consecutive MAXEIG analyses using 15, 30, 45, and 60 sliding windows. As noted earlier, increasing the number of windows brings the right-end peak of a low base-rate taxon into sharper definition. Such a change in curve shape did not occur with the BDI data (see Figure 2, top row) nor with the simulated dimensional data (middle row) but did occur with the simulated taxonic data (bottom row). Moreover, the base-rate estimate for the BDI data (.01) almost perfectly matched that of the simulated dimensional data (.02), whereas the base-rate estimate of the simulated taxonic data (.22) again resembled the simulated taxon base rate of .23.

L-Mode Analyses

L-Mode was performed using the 10 selected BDI items as indicators (see Waller & Meehl, 1998, for L-Mode program code). The resulting distribution of factor scores was unimodal (see Figure 3), and the three base-rate estimates derived from this procedure (.99, .30, and .46) were markedly divergent. The simulated dimensional data also yielded a unimodal distribution of factor scores with similarly discrepant base-rate estimates (.98, .32, and .40), whereas the simulated taxonic data yielded a bimodal distribution of factor scores with base-rate estimates (.27, .24, and .22) that agreed well with one another and with the simulated base rate of .23.

[4] Separate graphs for men and women are available on request from Ayelet Meron Ruscio.

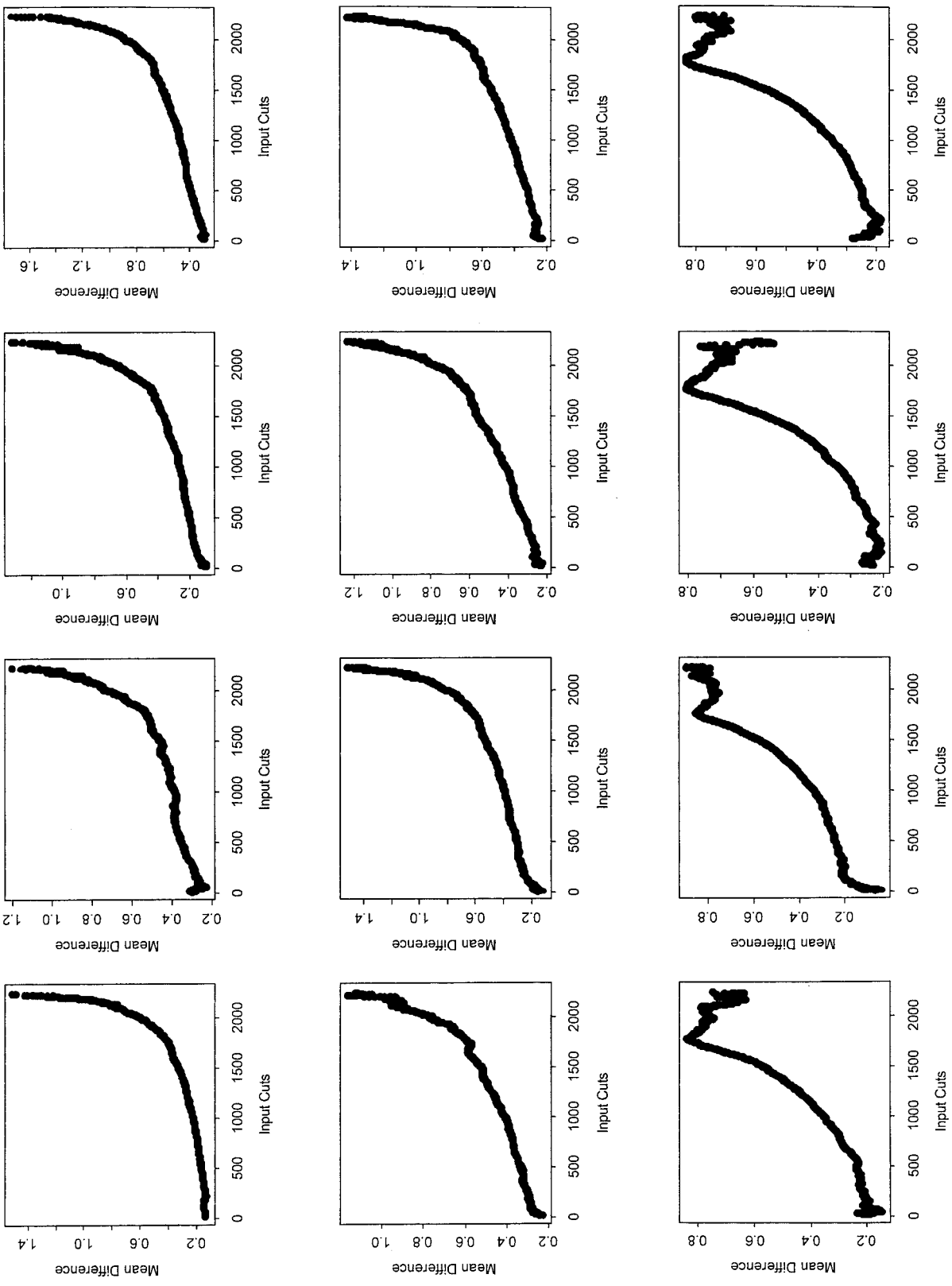

Figure 1. MAMBAC (mean above minus below a cut) curves, with cuts made at each case along the input indicator (*x*-axis) and the mean difference between those cases above and below the cut on the output indicator plotted on the *y*-axis. Top row: Curves generated from Beck Depression Inventory data. Middle row: Curves generated from simulated dimensional data. Bottom row: Curves generated from simulated taxonic data. To conserve space, we present only the first four curves from each series of 10 as a representative sample. The full panels are available on request from Ayelet Meron Ruscio.

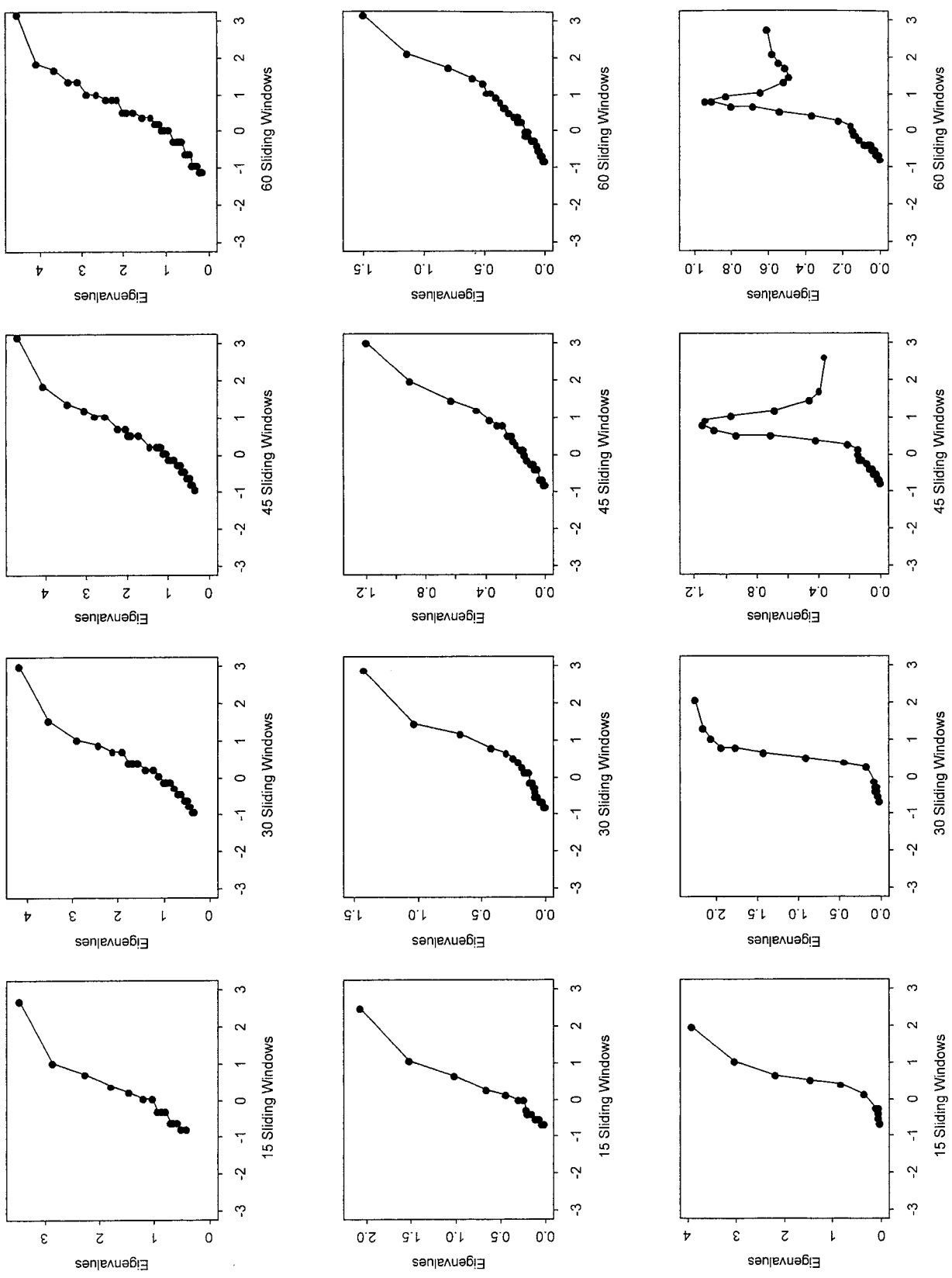

Figure 2. MAXEIG (maximum eigenvalue) curves plotting the eigenvalue of the indicator covariance matrix's first principal factor (*y*-axis) within sliding windows (either 15, 30, 45, or 60) along the input indicator (*x*-axis). Top row: Curves generated from Beck Depression Inventory data. Middle row: Curves generated from simulated dimensional data. Bottom row: Curves generated from simulated taxonic data.

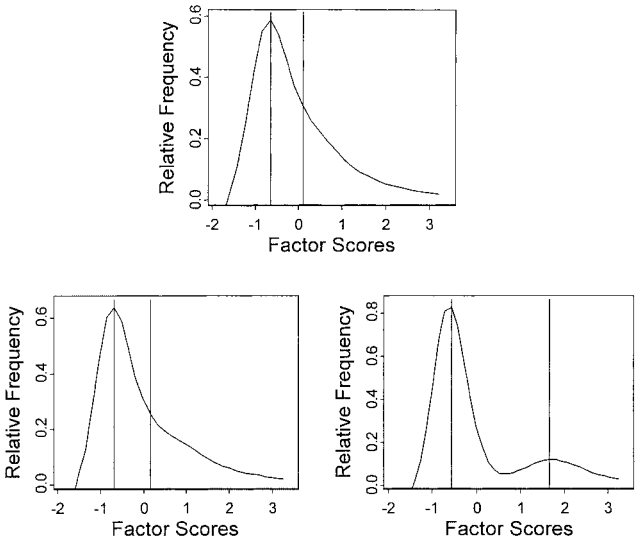

Figure 3. L-Mode (latent mode) curves depicting the distribution of scores on the first principal factor derived from a factor analysis of all 10 indicators. Top graph: Curve generated from Beck Depression Inventory data. Left graph: Curve generated from simulated dimensional data. Right graph: Curve generated from simulated taxonic data.

Discussion

The present investigation provides strong evidence for the latent dimensionality of BDI depression in analogue samples. BDI items encompassing a broad array of depressive symptoms were selected for their ability to validly distinguish cases at the most commonly used (10) and the most widely recommended (16) BDI cut scores (Kendall et al., 1987). Three independent taxometric procedures performed with these indicators yielded dimensional results, and base-rate estimates of the putative depression taxon evidenced poor agreement within and across procedures, suggesting that neither of these cut scores correspond to natural groups. Results obtained with BDI data were nearly identical to those obtained with simulated dimensional data and quite distinct from those obtained with simulated taxonic data. Taken together, these results indicate that the underlying structure of analogue depression as measured by the BDI is dimensional in nature.

The present results have important implications for the assessment of depression by the BDI in analogue samples. First, they suggest that analogue depression may be most validly represented by a continuous BDI score. Given the apparent absence of discrete groups underlying the BDI score distribution, it may be concluded that *any* cut made along this continuum constitutes an arbitrary designation that needlessly increases measurement error and compromises statistical power. Such spurious classification not only may weaken or mask important relationships among variables but may even alter them, thereby impeding theoretical understanding of depression phenomena (cf. Fraley & Waller, 1998). We have argued elsewhere (J. Ruscio & Ruscio, 2002) that the utility of psychological assessment is optimized when one's measurement model (i.e., dimensional scaling or classification model) is matched to the latent structure of the construct being assessed. Given the apparent dimensionality of BDI depression in our analogue sample, we suggest that BDI cut scores are inherently artificial and misleading in this population and should therefore be avoided. Similarly, it may be more appropriate to describe analogue cases falling at different levels of the BDI distribution by their depressive severity than by labels such as *depressed* or *dysphoric*, which connote the existence of a meaningful divide between those with and without such labels.

At the same time, despite the dimensional nature of the BDI in analogue populations, researchers may still choose to cut the BDI distribution for certain pragmatic purposes (e.g., selecting a group of highly depressed students for research) or simply because they prefer to work with categories rather than dimensions. In such cases, it will be important for researchers to clearly indicate that they are working with a dimensional construct, to provide a rationale for their decision to divide this dimension (weighed against the increased error and loss of power introduced by spurious categorization), to explain the suitability and utility of their chosen cut score relative to other possible cut scores within the population of interest, and to demonstrate that inferences drawn from the constructed groups are reasonable and appropriate, given the latent structure of analogue depression.

The present findings have related implications for investigations of analogue depression that use the BDI. More specifically, they suggest that comparative research designs, which rely on the construction of depressed and nondepressed groups for statistical comparisons, may be inappropriate for use with analogue participants. A more fruitful methodological approach would instead seek to identify correlates or predictors of the entire available range of depressive severity scores on the BDI (see A. M. Ruscio, Borkovec, & Ruscio, 2001). Such an approach would enhance the statistical power of investigations and allow researchers to search for nonlinear (e.g., quadratic, step function, logarithmic) relationships between BDI depression and measures of psychopathology (anxiety, substance abuse), emotion (positive and negative affect), cognition (rumination, information processing), behavior (social interaction, treatment seeking, suicidality), and other variables of interest. By virtue of their enhanced power to detect significant effects, as well as their ability to uncover relationships that differ at varying levels of depressive severity, such studies should lead to a deeper understanding of depression and related phenomena in analogue samples.

The present study sought to evaluate the validity of the common practice of dividing analogue samples into "depressed" and "nondepressed" groups on the basis of conventional cutoff scores on the BDI. To provide a fair and representative test of this practice, we used a large, mixed-gender, college student sample that was much like the analogue samples typically used in depression research. Moreover, we used the BDI because of its overwhelming popularity in analogue studies of depression, thus making it the measure whose underlying structure would be of greatest interest and relevance to depression researchers. However, although these findings strongly suggest that the latent structure of the BDI is dimensional in analogue samples, their ability to more broadly comment about the latent structure of depression is limited in two ways. First, our analyses were based solely on the BDI, and other self-report measures (e.g., Zung Self-Rating Depression Scale; Zung, 1965), interview rating scales (e.g., Hamilton Rating Scale for Depression; Hamilton, 1960), or structured clinical interviews with mood disorder modules (e.g., Structured Clinical Interview for *DSM–IV*; First, Spitzer, Gibbon, & Williams, 1996) may have yielded a different structural solution. Indeed, even the newly

published revision of the BDI (BDI–II; Beck, Steer, & Brown, 1996) may have produced different results. Although the high degree of overlap between the original and revised versions of the BDI ($r = .93$; Dozois et al., 1998) suggests that the two versions are likely to have essentially equivalent latent structure, it would be valuable to cross-validate the present results on BDI–II responses to determine this empirically. More generally, research is needed to test the latent structure of analogue depression using other measures of the construct, not only with college students but with community and other analogue samples as well.

Second, it is possible that a depression taxon existed in our analogue sample but was so small that our indicators (selected to be maximally valid at the BDI cut scores of 10 and 16) and taxometric procedures (whose ability to detect taxa with base rates lower than .10 is unknown) were unable to uncover it. It is important to emphasize that the goal of the present study was to ascertain whether the cut scores most commonly used and recommended for the BDI actually correspond to latent groups, not to determine whether a major depression taxon could be identified in this analogue sample. Thus, although our findings indicate that no qualitative boundaries exist at the BDI cut scores of 10 or 16, they do not rule out the possibility that a depression taxon—demarcated by a far higher BDI score—exists in analogue samples. Despite these important caveats, however, the convergence of these findings with those of taxometric analyses performed in clinical samples using multiple measures of depression (J. Ruscio & Ruscio, 2000) makes a taxonic solution for depression seem increasingly unlikely.

The present findings may also promote understanding of the controversial relationship between analogue depression and the depression experienced by clinical patients. Several investigators have questioned whether the BDI specifically assesses depression in analogue cases or whether it is instead a more diffuse measure of general symptomatology, distress, or negative affect in this population (e.g., Coyne, 1994; Gotlib, 1984). The present data do not bear directly on the ongoing debate about the comparability of analogue and clinical depression, as the sample under investigation was purely analogue in nature. However, the latent dimensionality of BDI depression in both the present analogue sample and in clinical samples (J. Ruscio & Ruscio, 2000) argues against the existence of a depression taxon at either extreme of the depressive severity continuum. This argument is consistent with recent findings of phenomenological continuity across the full range of depressive severity in community samples spanning adolescence to late adulthood (Lewinsohn, Solomon, Seeley, & Zeiss, 2000) and with a growing research literature revealing continuity between mild and severe forms of depressive experience (e.g., Flett et al., 1997; Judd et al., 1997; Kendler & Gardner, 1998). Thus, although research is needed to more directly elucidate the relationship between analogue and clinical depression, the present study may be seen to contribute at least indirectly to continuing dialogue on this issue.

Finally, the present study was the first to describe and demonstrate the effects of severe skew on the results of different taxometric procedures. Although prior Monte Carlo investigations (Cleland & Haslam, 1996; Haslam & Cleland, 1996) concluded that skew only modestly influences the shapes of taxometric curves, these investigations analyzed data sets that were simulated to have relatively low levels of skew. For this reason, we simulated a taxonic and a dimensional data set whose parameters (including indicator skew) matched those of our BDI data. Analyses of our BDI data revealed that strong positive skew affected the shapes of the taxometric curves in significant but predictable ways, making them appear more ambiguous and thus more difficult to interpret without a comparative benchmark. However, analysis of the simulated data sets revealed clear differences between taxonic and dimensional data even in the presence of considerable skew, and comparison of these results to those of our BDI data greatly facilitated curve interpretation. These findings underscore the value of simulating comparison data in taxometric research, both to assess whether the parameters of the available data are capable of distinguishing taxonic from dimensional structure and to aid in the evaluation of taxometric results (J. Ruscio, 2001). Also, as high levels of indicator skew are common in the investigation of low base-rate phenomena, these findings highlight the need for additional Monte Carlo research to identify the degree of skew that can be tolerated by each taxometric procedure under a variety of data and sample conditions.

The current study used the taxometric method to evaluate the latent structure of the BDI in a large analogue sample. Converging evidence suggested that analogue depression, as measured by the BDI, is dimensional in nature. These findings speak strongly against the common practice of classifying individuals into depressed or nondepressed groups on the basis of their BDI scores. Instead, they suggest that a shift toward the use of continuous BDI scores and correlation–regression analysis models will promote the most rapid and thorough accumulation of knowledge about depression phenomena.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andersen, S. M., Spielman, L. A., & Bargh, J. A. (1992). Future-event schemas and certainty about the future: Automaticity in depressives' future-event prediction. *Journal of Personality and Social Psychology, 63,* 711–723.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Beck, A. T., & Steer, R. A. (1993). *Beck Depression Inventory manual.* San Antonio, TX: Harcourt Brace.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory manual* (2nd ed.). 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.

Broadhead, W. E., Blazer, D. G., George, L. K., & Tse, C. K. (1990). Depression, disability days, and days lost from work in a prospective epidemiologic survey. *Journal of the American Medical Association, 264,* 2524–2528.

Cleland, C., & Haslam, N. (1996). Robustness of taxometric analysis with skewed indicators: I. A Monte Carlo study of the MAMBAC procedure. *Psychological Reports, 79,* 243–248.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7,* 249–253.

Compas, B. E., Ey, S., & Grant, K. E. (1993). Taxonomy, assessment, and diagnosis of depression during adolescence. *Psychological Bulletin, 114,* 323–344.

Coyne, J. C. (1994). Self-reported distress: Analogue or ersatz depression? *Psychological Bulletin, 116,* 29–45.

Deardorff, W. W., & Funabiki, D. (1985). A diagnostic caution in screen-

ing for depressed college students. *Cognitive Therapy and Research, 9,* 277–284.

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory—II. *Psychological Assessment, 10,* 83–89.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM–IV Axis I Disorders—Patient edition* (SCID–I/P, Version 2.0). Biometrics Research Department: New York State Psychiatric Institute, New York.

Flett, G. L. Vredenburg, K., & Krames, L. (1997). The continuity of depression in clinical and nonclinical samples. *Psychological Bulletin, 121,* 395–416.

Fraley, R. C., & Waller, N. G. (1998). Adult attachment patterns: A test of the typological model. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 77–114). New York: Guilford Press.

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

Grove, W. M., & Andreasen, N. C. (1989). Quantitative and qualitative distinctions between psychiatric disorders. In L. N. Robins & J. E. Barrett (Eds.), *The validity of psychiatric diagnoses* (pp. 127–139). New York: Raven Press.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.

Haslam, N., & Cleland, C. (1996). Robustness of taxometric analysis with skewed indicators: II. A Monte Carlo study of the MAXCOV procedure. *Psychological Reports, 79,* 1035–1039.

Horwath, E., Johnson, J., Klerman, G. L., & Weissman, M. M. (1992). Depressive symptoms as relative and attributable risk factors for first-onset major depression. *Archives of General Psychiatry, 49,* 817–823.

Ingram, R. E., & Hamilton, N. A. (1999). Evaluating precision in the social psychological assessment of depression: Methodological considerations, issues, and recommendations. *Journal of Social and Clinical Psychology, 18,* 160–180.

Johnson, J., Weissman, M. M., & Klerman, G. L. (1992). Service utilization and social morbidity associated with depressive symptoms in the community. *Journal of the American Medical Association, 267,* 1478–1483.

Judd, L. L., Akiskal, H. S., & Paulus, M. P. (1997). The role and clinical significance of subsyndromal depressive symptoms (SSD) in unipolar major depressive disorder. *Journal of Affective Disorders, 45,* 5–18.

Katz, R., Shaw, B. F., Vallis, M. T., & Kaiser, A. S. (1995). The assessment of severity and symptom patterns in depression. In E. E. Becklam & W. R. Leber (Eds.), *Handbook of depression* (2nd ed., pp. 61–85). New York: Guilford Press.

Kendall, P. C., & Flannery-Schroeder, E. C. (1995). Rigor, but not rigor mortis, in depression research. *Journal of Personality and Social Psychology, 68,* 892–894.

Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., & Ingram, R. E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy & Research, 11,* 289–299.

Kendler, K. S., & Gardner, C. O., Jr. (1998). Boundaries of major depression: An evaluation of *DSM–IV* criteria. *American Journal of Psychiatry, 155,* 172–177.

Kessler, R. C., Zhao, S., Blazer, D. G., & Swartz, M. (1997). Prevalence, correlates, and course of minor depression and major depression in the national comorbidity survey. *Journal of Affective Disorders, 45,* 19–30.

Lewinsohn, P. M., Solomon, A., Seeley, J. R., & Zeiss, A. (2000). Clinical implications of "subthreshold" depressive symptoms. *Journal of Abnormal Psychology, 109,* 345–351.

Lips, H. M., & Ng, M. (1985). Use of the Beck Depression Inventory with three non-clinical populations. *Canadian Journal of Behavioral Science, 18,* 62–74.

Marsh, K. L., & Weary, G. (1994). Severity of depression and responsiveness to attributional information. *Journal of Social and Clinical Psychology, 13,* 15–32.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Meehl, P. E. (1999). Clarifications about taxometric method. *Applied and Preventive Psychology, 8,* 165–174.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52,* 194–216.

Meehl, P. E., & Yonce, L. J. (1994). Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychological Reports, 74,* 1059–1274.

Meehl, P. E., & Yonce, L. J. (1996). Taxometric analysis: II. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV procedure). *Psychological Reports, 78,* 1091–1227.

Pelham, B. W. (1991). On the benefits of misery: Self-serving biases in the depressive self-content. *Journal of Personality and Social Psychology, 61,* 670–681.

Pinkley, R., Laprelle, J., Pyszczynski, T., & Greenberg, J. (1988). Depression and the self-serving search for consensus after success and failure. *Journal of Social and Clinical Psychology, 6,* 235–244.

Ruscio, A. M., Borkovec, T. D., & Ruscio, J. (2001). A taxometric investigation of the latent structure of worry. *Journal of Abnormal Psychology, 110,* 413–422.

Ruscio, J. (2001, June). *Simulating comparison data to help interpret the results of taxometric analyses.* Paper presented at the annual meeting of the American Psychological Society, Toronto, Ontario, Canada.

Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology, 109,* 473–487.

Ruscio, J., & Ruscio, A. M. (2002). A structure-based approach to psychological assessment: Matching measurement models to latent structure. *Assessment, 9,* 4–16.

Sherbourne, C. D., Wells, K. B., Hays, R. D., Rogers, W., Burnam, M. A., & Judd, L. L. (1994). Subthreshold depression and depressive disorder: Clinical characteristics of general medical and mental health specialty outpatients. *American Journal of Psychiatry, 151,* 1777–1784.

Tennen, H., Eberhardt, T. L., & Affleck, G. (1999). Depression research methodologies at the social–clinical interface: Still hazy after all these years. *Journal of Social and Clinical Psychology, 18,* 121–159.

Tennen, H., Hall, J. A., & Affleck, G. (1995). Depression research methodologies in the *Journal of Personality and Social Psychology*: A review and critique. *Journal of Personality and Social Psychology, 68,* 870–884.

Vredenburg, K., Flett, G. L., & Krames, L. (1993). Analogue versus clinical depression: A critical reappraisal. *Psychological Bulletin, 113,* 327–344.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua.* Thousand Oaks, CA: Sage.

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

Appendix

S+ Program Code for MAMBAC

```
MAMBAC <- function(x, y)
{
        N <- length(x)
        z <- vector("numeric",N)
        xyz <- data.frame(x, y, z)
        xyz <- xyz[sort.list(xyz[,1]),]
        xyz[,3] <- seq(1,N)
        d <- vector("numeric", N - 48)
        for (i in (25:(N - 24)))
           d[i - 24] <- mean(xyz[xyz[3] > i, 2]) - mean(xyz[xyz[3] <= i, 2])
        return(d)
}

MAMBAC.Each.Indicator <- function(data, Ind = 0)
{
        if (Ind == 0)
            Ind <- dim(data)[2]
        N <- dim(data)[1]
        MAMBAC.Values <- matrix(rep(0, (N - 48) * (Ind + 1)), ncol=(Ind + 1))
        Sums <- rep(0,N)
        for (i in 1:Ind)
           Sums <- Sums + data[,i]
        for (i in 1:Ind)
        {
            Input <- Sums - data[,i]
            Output <- data[,i]
            MAMBAC.Values[,i] <- MAMBAC(Input, Output)
        }
        P.est <- vector("numeric",Ind)
        for (i in 1:Ind)
            P.est[i] <- 1/((MAMBAC.Values[N - 48,i]/MAMBAC.Values[1,i]) + 1)
        print(cat("Base rate estimates for each curve:", "\n"))
        print(cat(round(P.est,3), "\n", "\n"))
        print(summary(P.est))
        print(cat("\n","Mean P: ",round(mean(P.est),3),"\n"))
        print(cat("  SD P: ",round(sqrt(var(P.est)),3),"\n"))
        MAMBAC.Values[,Ind + 1] <- seq(25,N - 24)
        frame()
        par(mfrow=c(3,4))
        par(pty="")
        for (i in 1:Ind)
        {
            plot(MAMBAC.Values[,Ind + 1], MAMBAC.Values[,i], type="p", col=1, pch=16,
                  xlab="Input Cuts", ylab=paste("Mean Difference"))
        }
        return(N)
}
```

Received May 10, 2001
Revision received December 20, 2001
Accepted January 22, 2002 ■