Continuity and Discontinuity Models of Bulimia Nervosa:  
A Taxometric Investigation

David H. Gleaves  
Texas A&M University and  
the Renfrew Foundation

Michael R. Lowe  
MCP Hahnemann University and  
the Renfrew Foundation

Alicia C. Snow and Bradley A. Green  
Texas A&M University

Kathleen P. Murphy-Eberenz  
MCP Hahnemann University and  
the Renfrew Foundation

Past research evaluating the continuity and discontinuity models of bulimia has produced inconclusive results. In the current study, we performed a taxometric analysis of bulimia nervosa using means above minus below a sliding cut and maximum covariance analysis with a sample of women diagnosed with bulimia nervosa (n = 201) or women college students (n = 412). Indicators were derived from the Bulimia Test-Revised and the Eating Attitudes Test-26, and both a mixed sample and the nonclinical sample were analyzed. With both taxometric methods and both mixed and nonclinical samples, results were consistently suggestive of a latent taxon for bulimia. These results challenge a dimensional model of bulimia nervosa.

Most psychological models of bulimia nervosa assume that the disorder differs in degree, rather than in kind, from lesser forms of eating disturbance (Drewnowski, Yee, Kurth, & Krahm, 1994; Mintz & Betz, 1988; Polivy & Herman, 1987; Scarano & Kalodner-Martin, 1994; Stice, Ziemba, Margolis, & Flick, 1996; Striegel-Moore, Silverstein, & Rodin, 1986). According to this viewpoint, there is a proportional relationship between the degree of eating disorder risk factors (e.g., body dissatisfaction, chronic dieting) exhibited by an individual and the extent of that individual’s bulimic symptoms (e.g., binge eating, purging). Although there are variations in the particulars of different theorists’ models, all of them suggest that disordered eating exists on a continuum; thus, these perspectives have been referred to collectively as the continuum model of bulimia nervosa (Lowe et al., 1996; Ruderman & Besbeas, 1992). The discontinuity perspective, on the other hand, suggests that bulimia differs qualitatively from lesser forms of weight and eating concerns. The discontinuity perspective is exemplified by the work of Hilde Bruch (1973), who suggested that bulimia develops in young women who were exposed to particular mothering influences as children.

Studies have investigated the continuity and discontinuity models of bulimia by examining the nature of the relationship between the extent of putative risk factors for bulimia and the extent of bulimic symptomatology. The risk factors examined can be divided into those that are specifically related to eating and weight concerns (e.g., chronic dieting) and those representing more general risk factors (e.g., personality traits). Evidence consistent with both the continuity and discontinuity models has been found for both types of risk factor. For the continuity model, evidence of a continuous increase in eating and weight concerns across unrestrained eaters, restrained eaters, and bulimic individuals was found by Rossiter, Wilson, and Goldstein (1989), Laessle, Tuschl, Waadt, and Pirke (1989), Stice et al. (1996), and Lowe et al. (1996). Similar support for continuity on measures of general psychopathology was reported by Stice et al. and Lowe et al. As for the discontinuity perspective, Lowe et al. found evidence for discontinuity on an eating-disorder-specific symptom (binge eating severity), and Ruderman and Besbeas (1992), Rossiter et al. (1989), and Laessle et al. (1989) all found evidence of discontinuity on various measures of psychopathology. In sum, the results of the empirical studies conducted to date have been unable to resolve the issue of whether bulimia nervosa differs quantitatively or qualitatively from more common forms of eating and weight concerns.

Above and apart from the diversity of findings on this question, however, is the issue of whether the research designs that have been used to investigate continuity are capable of definitively distinguishing between the continuity and discontinuity models of bulimia. The problem is that the indicators used to address the continuity–discontinuity issue could produce results consistent
with either model whether bulimia is in reality dimensional or
categorical in nature. As Meehl (1995) noted,

Psychologists who think there must always be clear-cut boundaries
are mixing the indicators with the latent taxon they indicate. The
distinction between qualitative and quantitative, or between a quan-
titative variable having a step function and one that behaves smoothly
even in the discriminating region of interest), can occur in all four
combinations. . . . Both the latent entity and its manifest indicators can
be either qualitative or quantitative, and all degrees of overlap be-
tween quantitative indicators' distributions can occur. (p. 268)

Thus, the appearance of gradual, linear increases in bulimic symp-
toms across various groups of non-dieters, dieters, and bulimic
individuals (Drewnowski et al., 1994; Laessle et al., 1989; Rossiter
et al., 1989; Stice et al., 1996) does not necessarily mean that
bulimia is dimensional in nature. Conversely, if groups of dieters
and non-dieters do not differ from one another but both differ from
bulimic individuals on measures of psychopathology (Laessle et
al., 1989; Rossiter et al., 1989; Ruderman & Besbeas, 1992) or on
binge eating (Low et al., 1996), it does not necessarily mean that
bulimia is categorical in nature.

Meehl has demonstrated that it is not the distribution of indi-
vidual indicators of a construct that is of critical significance;
rather, it is the nature of the relationship among indicators that can
best determine if a construct is dimensional or discrete. He and his
colleagues have described a group of related statistical procedures
that are designed to provide a more definitive answer to the
question of whether a given disorder represents a qualitatively
distinct class (a taxon) or an extreme endpoint on a continuously
distributed dimension (Meehl, 1995; Meehl & Golden, 1982;
Waller & Meehl, 1997). In brief, these procedures are based on the
position that the relationship between indicators for a latent class
(a taxon) will be different from the relationship between indicators
of a latent dimension. As Waller, Putnam, and Carlson (1996)
noted, "We require our putative taxon indicators to cohere in a
specifiable manner that is psychometrically distinguishable from
that produced by dimensional variables" (p. 303).

Most of these taxometric procedures examine the nature of the
covariance between valid indicators of a disorder. If there are three
indicators of a disorder, the covariance between any two of these
indicators at all possible levels of the third indicator can be
determined. If all possible combinations of indicators are assessed
and the distribution of resulting covariances is graphed, the form of
the distribution of covariances will generally assume one of two
characteristic shapes, depending on whether the disorder under
study is taxonic or dimensional in nature (Meehl, 1995). The
validity of this method has been established by repeated demon-
strations that phenomena that are known to be taxonic in nature
(e.g., medical diseases with known organic causes) show distinctly
different covariance distributions than phenomena known to be
nontaxonic (e.g., body weight or intelligence).

Waller et al. (1996) also pointed out that the expected correla-
tion between indicators of a taxon is zero in taxon members and in
nonmembers, but greater than zero when the two groups are
combined. For example, in the case of the taxon of meningitis, they
noted that individuals with meningitis tend to experience both neck
pain and high fever. Healthy individuals most frequently experi-
ence neither of these symptoms. When the correlation between
neck pain and fever is examined within groups of taxon members
and nonmembers (meningitis sufferers and non-sufferers), it will
be close to zero. However, the correlation between these symp-
toms in a mixed sample of healthy people and meningitis victims
will be significantly greater than zero.

Two of our studies have provided comparable data on bulimia.
Low et al. (1996) studied four groups of individuals (unrestrained
non-dieters, restrained non-dieters, current dieters, and bulimic
individuals) who differed systematically in their dieting intensity.
When we examined binge eating severity using a clinically mean-
ingful measure of binge eating, all of the nonbulimic control
groups showed very low levels of binge eating. Thus, the corre-
lation between two indicators of bulimia (dieting and binge eating)
was essentially zero among nonbulimic individuals. Lowe,
Gleaves, and Murphy-Eberenz (1998) then studied the relationship
between dieting intensity and binge eating severity among bulimic
individuals (i.e., among those with the putative taxon). When
frequent and infrequent dieters were compared on binge eating
severity, they did not differ in one sample, whereas frequent
dieters were found to binge less than infrequent dieters in two
other samples. Overall, then, no (positive) correlation was found
between dieting intensity and binge severity in either individuals
without (Low et al., 1996) or those with (Low et al., 1998)
bulimia. In mixed groups of those with and without bulimia, on the
other hand, there is a correlation between binge eating and dieting
(Low et al., 1996; Ruderman & Besbeas, 1992). Although this
evidence is consistent with the assumption that bulimia is taxonic
in nature, more definitive information concerning the taxonicity of
bulimia requires the implementation of taxometric procedures spe-
cifically designed for this purpose. The goal of this study is to
conduct such a taxometric analysis.

The results of taxometric analyses are most interpretable when
the base rate of the conjectured taxon is around .50 (Waller et al.,
1996). Thus, in examining the taxometric status of a disorder, it is
useful to compare a disordered group with a roughly equal size
comparison group that has the potential for exhibiting some of the
indicators being investigated but are unlikely to possess the taxon
if it exists. For the disorder of bulimia, a good comparison group
would be female college students, because it is just this group that
has provided the evidence in support of a continuity model of
bulimia (Drewnowski et al., 1994; Herman & Polivy, 1984; Mintz
& Betz, 1988; Stice et al., 1996). Female college students are
young, typically from middle- to upper-middle-class backgrounds,
and are likely to be exposed to psychosocial influences that em-
phasize attractiveness, heterosexual relationships, and conformity
with prevailing values (Striegel-Moore, 1993). Thus, if bulimic
behavior exists on a continuum, female college students should
constitute an ideal reference group for comparison with bulimic
individuals.

Participants

A total of 613 women participated in the study. Noneclinical college
students (n = 412) were undergraduate students in an introduction
to psychology course at the southwestern university. They were offered course
credit for their participation. For the clinical bulimia nervosa group (n =
201), we used a data set (some of which was described by Low et al.,
1998) from a residential treatment facility for women with eating disorders.
All of the women were diagnosed as having bulimia nervosa on the basis
of a semistructured intake interview using Diagnostic and Statistical Manual of Mental Disorders, 3rd edition, revised (DSM-III-R; American Psychiatric Association [APA], 1987) criteria.

Assessment Measures

As a source of quantitative indicators for our taxometric analyses, we used the Bulimia Test—Revised (BULT-R; Thelen, Farmer, Wonderlich, & Smith, 1991) and the Eating Attitudes Test—26 (EAT-26; Garner, Olmsted, Bohr, & Garfinkel, 1982). The BULT-R is a 36-item self-report measure designed to assess eating behaviors and attitudes related to bulimia. Thelen et al. reported good reliability and validity of the instrument. As described below, we used both individual item scores as well as aggregate factor scores. The EAT-26 was originally a 40-item self-report inventory designed to assess thoughts and behaviors related to anorexia. The 26-item version was developed from factor analyses when items not loading on any factor were eliminated. The EAT-26 has been found to have good internal consistency and validity (Williamson, Anderson, & Geaves, 1996).

The choice of the BULT-R was fairly obvious, as it was designed to measure the behaviors characteristic of bulimia nervosa. We chose to also include the EAT-26 because it has more items measuring dieting behavior, which, research has suggested, is an important component of multidimensional models of bulimia (Geaves & Eberenz, 1995; Geaves, Williamson, & Barker, 1993; Tobin, Steinberg, Staats, Johnson, & Dennis, 1991).

Statistical Analyses

Although several taxometric procedures have been developed (see Waller & Meehl, 1997), the procedures that we used were the two described by Meehl (1995): means above minus below a cut (MAMBAC) and maximum covariance (MAXCOV).

MAMBAC. MAMBAC, described in detail by Meehl and Yonce (1994), involves the use of two indicators that are uncorrelated within groups and are considered to be valid indicators of a conjectured taxon. This procedure is based on the effect of sorting a sample on an indicator that discriminates the conjectured taxon and nontaxon groups. Suppose two groups, t and c (for taxon group and complement group), are different with respect to indicator z, so that the distributions may overlap but have different means. If all cases are perfectly sorted into taxon and nontaxon groups, then the mean difference between the groups (z_bar_t - z_bar_c) will be equal to z_bar_t - z_bar_c. If miscategorizations occur and some taxon and nontaxon members are randomly sorted into the wrong category, the new z_bar_t will be pulled down and the new z_bar_c will be pulled up. This effect decreases the group difference (z_bar_t - z_bar_c). The more misclassifications that occur, the smaller the difference between the means. Now, also assume that another variable x (uncorrelated with z within groups) also distinguishes the taxon from the nontaxon (complement) group. If z scores are plotted with x scores as the input variable, then choosing a cut score on x also determines the degree of scrambling of taxon and nontaxon groups. The cut divides the scores on z in a way that allows one to subtract the mean of z scores below the x cut from the mean of z scores above the cut to arrive at the difference score z_bar_t - z_bar_c. This can be repeated while moving the x cut from the left to right to arrive at various estimates of z_bar_t - z_bar_c. A graph of the difference scores yields a characteristic convex shape if the phenomenon under study is taxonic (Meehl, 1995). This characteristic shape is displayed in the top three panels of Figure 1. In the case of a latent dimension, the MAMBAC graph will have a characteristic concave shape, as depicted in the bottom three panels of Figure 1. In the current investigation, the MAMBAC analyses were performed using the S-Plus statistical package (Version 3.3; Statistical Sciences, Inc., 1995) and a program adapted from Meehl and Yonce (1994). As these authors had recommended, we used the lowess smoothing procedure.

MAXCOV. The second procedure, MAXCOV, uses three indicators of a construct and is the specific application of the general principle of taxometric procedures described above. That is, if the assumption that indicators y and z are independent within the conjectured taxon and non-taxon groups is true, then the correlation between y and z will be zero for individuals who have both very low and very high scores on x. This occurs because high and low scorers on x will primarily be taxon and nontaxon members, and indicators are not expected to correlate within (taxon and complement) groups. Assuming that all three indicators are valid indicators of a taxon, plotting the covariance of y and z across x intervals should yield a characteristic shape if the situation is taxonic. This characteristic shape would begin with no correlation in the first x interval (nontaxon group), followed by ever-increasing correlations in successive x intervals until the x interval represents an even mix of taxon and nontaxon members for which there is a maximum correlation (hitmax interval) between y and z. Continuing up the x intervals results in greater and greater proportions of taxon members yielding an ever-decreasing correlation that eventually reaches zero when the group is composed only of taxon members. If the situation is truly dimensional, and individuals have been sampled from a single population, then the covariance between y and z will be randomly distributed around an average value. The resulting plot will be basically a flat line. Figure 2 illustrates groups of taxonic (top) and dimensional (bottom) plots generated by MAXCOV. More detailed discussion of this procedure can be found in Meehl and Golden (1982) and Meehl and Yonce (1996). In the current investigation, MAXCOV procedures were also carried out using the S-Plus statistical package. The MAXCOV program was adapted from one written by Waller (N. G. Waller, personal communication, January 7, 1998; also now in Waller & Meehl, 1997). As with the MAMBAC procedure, we used the lowess smoothing procedure.

Base rate estimation. With both MAMBAC and MAXCOV, it is possible to estimate the taxon base rate. If these estimates are found to be relatively consistent across multiple applications of the procedure(s), this is thought to be additional evidence that the construct is indeed taxonic. If base rate estimates vary widely, this can be construed as prima facie evidence that the taxonomic model is unsupported (Waller & Ross, 1997).

Procedures

With MAMBAC and MAXCOV, as well as with other taxometric procedures, it is recommended that conclusions regarding dimensionality or taxonicity not be based on single analyses (e.g., Waller & Meehl, 1997). Rather, a multiple consistency, multiple indicators approach is typically used. It is for this reason that we used more than one taxometric method with more than one approach to generating indicators. For all taxometric procedures, we conducted several general sets of analyses. First, we used a combined sample of clinical bulimia nervosa patients and nonclinical college students and performed MAMBAC and MAXCOV analyses using both empirically and theoretically derived indicators (described below). For the analysis with the mixed sample, we specifically included an approximately equal number of persons with and without a diagnosis of bulimia. Persons in the nonclinical group who scored 104 or higher on the BULT-R (the cutoff used by Thelen et al., 1991) were included in the bulimia nervosa group. This addition of 8 participants resulted in the first set of analyses being conducted with 418 participants, 209 in each group. We chose this 50/50 mixture because the taxometric procedures that we used are most interpretable when the base rate of the alleged taxon is .5 (Meehl & Yonce, 1994, 1996), and we reasoned that, if

1 Both instruments were scored using the recommendations of the original authors. The five options of each BULT-R item are simply scored 1–5 (or the reverse). However, the six options of each EAT-26 item are scored 0, 0, 0, 1, 2, 3 (or the reverse). The distributional properties of indicators may affect the results of at least some taxometric procedures. Future taxometric research should investigate the effects of alternative scoring procedures on overall results.
bulimia nervosa is taxonic, then persons diagnosed as having the disorder would be likely to possess the taxon and those not diagnosed would be unlikely to possess it. The second set of analyses used just the sample of college students (n = 412), without eliminating the 8 with high scores on the BULIT-R. We performed this set of analyses for two purposes. The first was to be sure that whatever results we obtained with the mixed sample would be replicated in the nonclinical sample (i.e., make sure that the results were not an artifact of simply having two distinct groups as might occur in some circumstances; Meehl, 1992). The second purpose was to attempt to measure the base rate of the conjectured taxon in the college population, if it was found to exist.

Preliminary Indicator Selection

For the empirical approach to indicator selection, we followed recommendations of Meehl and Yonce (1994) and first performed three preliminary analyses, conducted using SPSSPC v.5.1. These were (a) identifying which items of the BULIT-R and EAT-26 best discriminated the clinical and nonclinical groups (using a 50/50 split of bulimics and nonbulimics), (b) testing to see which items had high item-total correlations with the BULIT-R (a global measure of bulimia nervosa) and eliminating those with poor correlations, and (c) testing for and eliminating within group nuisance covariance (sizeable correlations within the clinical group).  

One could argue that by using indicators with large-group separation but low within-group correlations, one is in effect stacking the deck in favor of taxonicity. That is, such procedures assume that indicators for taxonic constructs will be uncorrelated within groups of taxon or non-taxon members but highly correlated in mixed samples. We believe that there are two ways to address this concern. The first is that one would only be able to find indicators that meet these criteria in cases in which there is some evidence of latent taxonicity. Thus, being able to find items that meet these criteria should be regarded as an initial hurdle in a series of taxometric analyses. The second way to address the concern is by pointing out that we also used a second procedure to indicator selection, purely on the basis of theory and item content rather than by looking at between- or within-group correlations. Doing so allowed us to look for consistency across different methods of selecting indicators.

Nuisance covariation would generally refer to sizeable correlation within either the conjectured taxon or complement group. If bulimia nervosa is taxonic, then persons diagnosed as having the disorder would be likely to possess the taxon and those not diagnosed would be unlikely to possess it.
Because indicators for a conjectured taxon should be able to clearly discriminate those with and without the taxon, Meehl and Yonce (1994) recommended ideally looking for items with 2 SD differences between groups. With our data, no items met this relatively strict criterion, so we then tested for 1.5 SD separation on items. A total of 9 items from the BULIT-R and 3 from the EAT-26 that passed this test are presented in Table 1. Rather than reduce our criterion to a lower degree of separation, we decided to proceed with these 12 items for subsequent analyses. We then tested to see which items had high item-total correlations with the BULIT-R. All of the items passed this test in that all had item-total correlations of greater than .70. Finally, we checked for within-group covariance. BULIT-R Item 28 was highly correlated with EAT-26 Item 18 ($r = .76$); the latter item was also correlated with EAT-26 Items 3 ($r = .57$) and 21 ($r = .70$). There were also several high interitem correlations in the range of .50 to .60. Thus, to reduce nuisance covariance, we aggregated highly correlated items (greater than $r = .50$). Specifically, we aggregated BULIT-R Items 5 and 25 to make one aggregate variable, BULIT-R Item 28 and all three EAT-26 items to make a second aggregate, and BULIT-R Items 32 and 34 to make a third. This procedure left us with a total of five indicators: BULIT-R Items 15 and 17, and three aggregate variables. Although this step did not completely eliminate nuisance covariance, it was substantially reduced (the highest correlation was $r = .44$, and the rest were below $r = .40$).

For the theoretical approach to indicator selection, we first developed a list of what the indicators for bulimia nervosa should be on the basis of accepted descriptions of the disorder (e.g., APA, 1987) and previous research aimed at identifying the basic features of the disorder (Gleaves & Eberenz, 1995; Gleaves et al., 1993; Tobin et al., 1991). We then chose items from the BULIT-R and EAT-26 that, on the basis of content, seemed to best measure these constructs. The theoretical factors that we were attempted measure were (a) binging behavior, (b) purging behavior, (c) body dissatisfaction-concerns, and (d) restrictive behaviors (dieting and using other methods to lose weight). Items that were chosen to measure these features are shown in Table 2. Although the above factor-analytic studies also described a feature characterized by negative affect or personality disturbance, or both, such a disturbance is not a diagnostic feature of the disorder (APA, 1987), and we did not attempt to measure this dimension in the current investigation. Future taxometric research should attempt to do so.

---

4 We do not think the nuisance covariance in our sample led to any misinterpretation of dimensional plots as taxonic. According to Monte Carlo studies by Meehl and Yonce (1994), "Increasing nuisance covariance flattens the MAMBAC curves progressively" (p. 1083). Thus, higher nuisance covariance makes taxonic plots look progressively more like dimensional plots, not the reverse.
Table 1
Items From the BULIT-R and EAT-26 With 1.5 SD Discrimination of Clinical and Nonclinical Participants

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>When I feel that my eating behavior is out of control, I try to take rather extreme measures to get back on course (strict dieting, fasting, laxatives, diuretics, self-induced vomiting, or vigorous exercise).</td>
</tr>
<tr>
<td>9.</td>
<td>How long have you been binge eating (eating uncontrollably to the point of stuffing yourself)?</td>
</tr>
<tr>
<td>10.</td>
<td>How often do you intentionally vomit after eating?</td>
</tr>
<tr>
<td>17.</td>
<td>After I binge eat, I turn to one of several strict methods to try to keep from gaining weight (vigorous exercise, strict dieting, fasting, self-induced vomiting, laxatives, or diuretics).</td>
</tr>
</tbody>
</table>

Table 2
Items From the BULIT-R and EAT-26 Chosen to Represent Theoretical Factors of Bulimia Nervosa

<table>
<thead>
<tr>
<th>Factor</th>
<th>BULIT-R items</th>
<th>EAT-26 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binging behavior</td>
<td>2, 3, 8, 9, 10, 13, 21, 22, 23, 30, 34</td>
<td>3, 4</td>
</tr>
<tr>
<td>Purgging behavior</td>
<td>5, 6, 15, 18</td>
<td>9, 26</td>
</tr>
<tr>
<td>Body concerns</td>
<td>4, 7, 12, 14, 24, 32, 35</td>
<td>1, 10, 11</td>
</tr>
<tr>
<td>Restrictive behaviors</td>
<td>11, 19, 20, 25, 29</td>
<td>6, 7, 12, 16, 17, 23</td>
</tr>
</tbody>
</table>

Reliability data are presented in Table 3. Across all of the plots rated, the two raters agreed 90.7% of the time ($\kappa = .67$). There was higher agreement with the mixed sample in which raters agreed on 100% of the plots. When there were disagreements between raters, it happened most frequently when one rater had rated a plot as ambiguous and the other rated it as either taxonic or dimensional. Of all the plots, only one was rated as taxonic by one rater and dimensional by the other. This was one of the MAXCOV plots with the nonclinical sample.

Results Based on Mixed Sample

MAMBAC results. As noted above, we began by analyzing data for empirically derived indicators using a 50/50 ratio of bulimia nervosa and non-bulimia nervosa participants ($n = 418$). Each item was paired with all other items. Reverse-ordered plots were also generated for each item pair (i.e., 1,2 is different from 2,1). With the five empirically derived indicators, this procedure led to a total of 20 item-level MAMBAC plots ($5!/(5-2)! = 20$). Examination of the plots led to a clear and consistent conclusion. All of the plots had the convex shape characteristic of taxonic data, and all of the plots were rated as taxonic by both raters. The average estimated base rate was .51 ($SD = .04$).

Using the four theoretically derived indicators, we were able to generate 12 additional MAMBAC plots. As with the empirically derived indicators, none of these plots demonstrated the concave shape that would be generated from dimensional data, and all were rated as being a taxonic plot. Representative plots from these analyses are presented in the top of Figure 3. The average base rate estimate was .49 ($SD = .07$).

MAXCOV results. With the five empirically derived indicators, there were 30 possible combinations of three indicators ($5!/$
both MAMBAC and MAXCOV analyses, we predicted a bulimic in the previous analysis. We reasoned that if bulimia nervosa is truly taxonic, then two patterns should emerge in subsequent analyses. First, the procedures would again give evidence of taxonicity, and second, the shapes of the plots would resemble a low base rate taxonic plot. 7 Because it was likely that some of the indicators would have skewed distributions in the nonclinical sample, we wanted to determine if any of our results were possibly artificial. We, thus, examined the skew of our data.

The results of these MAXCOV analyses were again quite consistent and fairly striking. A group of representative plots is shown in the bottom of Figure 3, and all of these (in Figure 3) were rated by both raters as being indicative of taxonicity. As can be seen, each has the characteristic convex shape suggestive of a latent taxon with a base rate somewhere close to .50. With our four theoretically derived indicators, we used the above described procedure of taking all possible pairs of indicators and then aggregated the remaining three indicators into a single variable across which the covariance of the first two variables was plotted. This procedure allowed us to produce 10 MAXCOV plots ((5!/(2!)) five pairs). 5

The results of these MAXCOV analyses were again quite consistent and fairly striking. A group of representative plots is shown in the bottom of Figure 3, and all of these (in Figure 3) were rated by both raters as being indicative of taxonicity. As can be seen, each has the characteristic convex shape suggestive of a latent taxon with a base rate somewhere close to .50. With our four theoretically derived indicators, we used the above described procedure of taking all possible pairs of indicators and then aggregated the remaining three indicators into a single variable across which the covariance of the first two variables was plotted. This procedure allowed us to produce 10 MAXCOV plots ((5!/(2!)) five pairs). 5

The results of these MAXCOV analyses were again quite consistent and fairly striking. A group of representative plots is shown in the bottom of Figure 3, and all of these (in Figure 3) were rated by both raters as being indicative of taxonicity. As can be seen, each has the characteristic convex shape suggestive of a latent taxon with a base rate somewhere close to .50. With our four theoretically derived indicators, we used the above described procedure of taking all possible pairs of indicators and then aggregated the remaining three indicators into a single variable across which the covariance of the first two variables was plotted. This procedure allowed us to produce 10 MAXCOV plots ((5!/(2!)) five pairs). 5

Table 3

<table>
<thead>
<tr>
<th>Procedure/sample composition</th>
<th>No. of plots rated</th>
<th>Agreement</th>
<th>( \kappa )</th>
<th>Rater 1</th>
<th>Rater 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAMBAC</td>
<td>Mixed</td>
<td>32</td>
<td>1.0</td>
<td>1.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Nonclinical</td>
<td>64</td>
<td>.89</td>
<td>.47</td>
<td>84.4</td>
</tr>
<tr>
<td></td>
<td>Simulated</td>
<td>12</td>
<td>.67</td>
<td>.36(^b)</td>
<td>00.0</td>
</tr>
<tr>
<td>MaxCOV</td>
<td>Mixed</td>
<td>16</td>
<td>1.0</td>
<td>1.0</td>
<td>93.8</td>
</tr>
<tr>
<td></td>
<td>Nonclinical</td>
<td>16</td>
<td>.88</td>
<td>.52</td>
<td>68.8</td>
</tr>
</tbody>
</table>

Note. MAMBAC = mean above minus below a cut; MAXCOV = maximum covariance.

\(^a\) This number is twice that for the mixed sample because it includes plots with transformed variables. \(^b\) The disagreements here concerned discriminating dimensional from "ambiguous" plots. None were rated as taxonic; thus, agreement as to whether plots were taxonic or not taxonic was 100%.

Analyses With College Sample

MAMBAC results. With the college sample \((n = 412)\) and the empirically derived indicators, we were again able to plot 20 MAMBACs. The results were once again quite consistent. In most cases, the peak was shifted to the right, as would be expected if there were a taxon base rate of .10 or less (Meehl & Yonce, 1994). The average estimated base rate was .15 \((SD = .15)\). With the theoretically derived indicators, we again plotted 12 MAMBACs. In all 11 out of 12 appeared to be indicative of taxonicity. The only plot that looked more like what would be expected with a dimensional construct (and was rated as such by both raters) was "body concerns" against "restrictive behaviors." However, the same two variables plotted in the reverse order produced a plot indicative of taxonicity. The mean base rate estimate (not including the graph with the dimensional shape) was .25 \((SD = .08)\).

Waller (personal communication, March 25, 1998) noted that there are some situations in which MAMBAC plots may look somewhat taxonic even when the latent construct is dimensional. Specifically, if data are highly positively skewed, the plots may resemble a low base rate taxonic plot. Because it was likely that some of the indicators would have skewed distributions in the nonclinical sample, we wanted to determine if any of our results were possibly artificial. We, thus, examined the skew of our data.

5 Note that this is one half the number of MAMBAC pairs, because the covariance of variables 1, 2 is the same as 2, 1. With the MAMBAC procedure, a plot of 1, 2 is different from 2, 1.

6 Finding base rate estimates of close to .50 in and of itself is not a convincing sign of taxonicity, as dimensional constructs will also produce base rate estimates of close to .50 (Meehl & Yonce, 1994, 1996), because they are basically random numbers from 0.0 to 1.0. However, the range of base rate estimates in the current study was very similar to those produced in Monte Carlo simulations with taxonic data with characteristics similar to ours (e.g., .15 SD separation on indicators, sample size 300–600; see Meehl & Yonce, 1996, pp. 1190–1191). A more important finding would be obtaining base rate estimates consistently greater or less than .50 on average when predicted by theory. This was the finding with the nonclinical sample.

7 It is possible that extreme platykurtosis could cause the same problem. Our method of simulating data also addressed the issue of kurtosis.
indicators and, where it was a problem, transformed the data. All of the indicators were positively skewed to some degree with the largest amounts of skew for the variables related to bingeing and purging. We then attempted square-root or log transformations, or both, to reduce the skew. We were successful in reducing the skew to less than 1.0 for all of the variables except BULIT-R Item 15 ("How often do you intentionally vomit after eating?") because the vast majority of the participants endorsed the lowest response. We then re-ran all of the MAMBAC analyses with the transformed variables.

The transformation of the variables led to an interpretative change in only one of the plots. We noted above that, with the untransformed data, the plot of body concerns against restrictive behaviors looked dimensional but that the reverse plot still looked taxonic. With the transformed variables, both of these plots now looked dimensional. Representative plots are depicted in the top panels of Figure 4. Average base rate estimates for the transformed variables (not including the dimensional plots) were .19 (SD = .18) for the empirical indicators and .24 (SD = .06) for the theoretical indicators.

Because we were unable to completely normalize all of the variables (i.e., a skew of 1.0 is still moderately skewed), we ran one more set of analyses to determine if our results were possibly an artifact of skewed data. We performed a Monte Carlo analysis by generating a dimensional data set with the distributional characteristics (mean, skew, kurtosis, and interitem correlations) of our own transformed data. To generate this data set, we used the PRELIS 2 program (Jöreskog & Sörbom, 1996) and the formulas from Fleishman (1978) for simulating non-normal distributions. We then ran MAMBAC on the dimensional data set. Representative plots are depicted in the bottom panels of Figure 4. Most of these still looked dimensional, even though the peak on the right was higher than on the left (apparently because of the skew). Two plots did superficially resemble a low base rate taxonic plot (e.g., the middle plot). However, close examination suggested a dissimilarity with those presented in Meehl and Yonce (1994) and with...
most of the plots generated from our data. Basically, these plots from simulated data looked like one half of a dimensional plot (which is what they are) rather than a taxonic plot with the peak shifted to the right. Thus, our finding of taxonic-looking MAMBAC plots did not seem to be due purely to the distribution of the indicators.

**MAXCOV results.** With the MAXCOV procedure and the empirically derived indicators, we were again able to produce 10 plots, using the procedure described above. Although the plots were not as clear as we had seen with the 50/50 split, there was a fairly consistent pattern. Most plots either demonstrated a peak toward the right boundary or were characterized by a steep incline with a peak at the right boundary. Because the MAXCOV program is unable to estimate the taxon base rate when the peak is at the boundary of the plot, base rate estimates were only possible in seven analyses. Thus, we could not calculate an unbiased average. The mean of the seven estimates was .27 (SD = .19) with a median of .20.

Similar results were found when we used the theoretically derived indicators to create six MAXCOV plots using all possible pairs and aggregating the remaining variables. Plots again generally either peaked at the right boundary or were convex with the peak toward the right boundary. Representative plots are depicted in Figure 5. Base rate estimates could be calculated in only three of the six analyses. These were .24, .22, and .33 (M = .26, Mdn = .24, SD = .06).

**Discussion**

Although several studies have been conducted on the continuity versus discontinuity models of bulimia, this was the first study to use methods specifically designed to distinguish types from continua.

---

It could be argued that the failure to estimate base rates in all analyses was problematic and indicative of a pseudotaxon rather than a taxon. We do not believe this to be the case. In each of the situations in which MAXCOV was unable to estimate a base rate, the graph peaked at the right boundary, which simply suggests a low base rate situation. The problem did not occur with MAMBAC because base rate estimates are calculated differently; however, the estimates from the two procedures are quite consistent. Taken together, these data suggest that we detected a real taxon with a moderately low base rate.
The vast majority of the results of this study were inconsistent with a strictly dimensional model of bulimia nervosa. That is, the results suggested the presence of a latent taxon or a latent discontinuity. The possible detection of a latent taxon does not, however, imply that there is not one or more dimensions also related to bulimia.

One of the principles of taxometric research is that multiple consistency tests—multiple hurdles need to be used. That is, the results of any single taxometric analysis are inconclusive. Only when consistency can be demonstrated can one make statements about taxonicity versus dimensionality. In the current study, we implemented such a strategy by using two different taxometric procedures with both empirically and theoretically derived indicators and by using both a mixed and a nonclinical sample. With one exception, we consistently found results that were suggestive of a latent taxon for bulimia. The one exception was the MAMBAC plots of dieting behavior and body concerns, which will be discussed below.

In the current study, the MAMBAC plots did seem to be affected, to some degree, by the distributional properties of the sample (i.e., skewness or kurtosis, or both). However, plots from dimensional data with similar distributional properties were fairly distinguishable from our eating disorder data. Future research should, however, examine the effects of using taxometric procedures with varieties of nonnormal data, and researchers using these procedures should interpret their results cautiously if they are based on such data.

Also, regarding the use of taxometric methods, we believe that this is the first study to report interrater reliability of taxometric plot ratings. We were encouraged that two raters produced such highly reliable data. The results of this study along with those of Snow (1997) suggest that taxometric procedures can be interpreted in a reliable fashion. In that study, an unpublished dissertation, two raters agreed on 148 (97%) of 152 plots (κ = .95). Base rate estimates were also fairly consistent. With the mixed sample, across both methods and both types of indicators, the average base rate estimate was .50, which is approximately what would be expected with a 50/50 sample. With the nonclinical sample, but without excluding potentially bulimic participants, the average base rate estimate was .23. This seems somewhat high, but there is evidence that the estimates from both MAMBAC and MAXCOV were probably positively biased. Meehl and Yonce (1994) noted that, with low base rate situations, the MAMBAC program may yield a positively biased estimate of the base. With the MAXCOV analyses, several base rates could not be estimated because the plots peaked at the right boundary; thus, the actual
estimate for these particular analyses would have been closer to zero and would have pulled down the average. The important point for the current investigation is that the base rate estimates were consistently below .50 (which is what would be obtained in a nontaxonic situation, in which they are basically random), which continues to suggest the presence of a latent taxon. Future research will be needed to more accurately measure the base rate of the conjectured bulimia nervosa taxon.

If we have identified a latent taxon for bulimia nervosa, then one of the next important questions will be What exactly is this taxon? This question cannot be conclusively answered from the current results. It is possible that the taxon (or a taxon in general) represents a genetic or otherwise biological vulnerability. However, environmental events may also be partially or completely responsible. Meehl (1992) also described what he referred to as "environmental mold taxa" (p. 147), which emerge when "persons subjected to certain (formal or informal) learning experiences—precepts, models, and reinforcement schedules—acquire motives, cathexes, cognitions, and act dispositions" (p. 149). It is also possible that a taxon can be due to an interaction of genetic and environmental influences, for example, phenylketonuria, which requires a particular gene and diet (phenylalanine) before pathology is expressed.

Although we do not believe that the current investigation can thoroughly answer the question of what exactly the taxon of bulimia nervosa would be, we consider this study a starting point. The variables found to be indicative of taxonicity are the best place to start in terms of defining the taxon. The empirically derived indicators (listed in Table 1) all seem to measure bingeing behavior, purging behavior, and food preoccupation. None of the items measuring dieting behavior and body dissatisfaction—concerns were included here. Furthermore, when we examined the theoretically derived indicators, and the indicators for these latter two constructs were plotted against one another, the result was a dimensional plot. Thus, the "best" indicators for the taxon, from the possible indicators that we used, may be bingeing, purging behavior, and food preoccupation. Dieting and body dissatisfaction in the absence of any other indicators do not appear to be indicative of the latent taxon. This is perhaps consistent with the findings of Joiner, Wonderlich, Salatsky, and Schmidt (1995), who found body dissatisfaction to be more associated with depression than bulimia. However, in our study, when body dissatisfaction was coupled with any of the other indicators (bingeing, purging, food preoccupation), there was evidence of taxonicity. Dieting and body dissatisfaction, taken alone or in combination, were not necessarily indicators of bulimia, but taken in conjunction with any of the other three signs were indicative of bulimia. Future research needs to examine what other variables may function as indicators for a bulimia nervosa taxon, including the variable of affective and personality disturbance described above.

If bulimia nervosa is indeed taxonic then another important question is What difference does it make? (Meehl, 1992). That is, why should one care if bulimia nervosa is taxonic or dimensional; does it, or should it matter? Meehl (1992) described five reasons regarding taxometric research in general, only three of which we discuss here in the context of bulimia. The first reason may be regarded by some as purely academic, but as Meehl (1992) put it, "if there are real taxa in a domain, theoretical science should come to know them" (p. 161). That is, if the goal of science is, as Plato wrote, to "carve nature at its joints" (cited in Gangestad & Snyder, 1985), then we as scientists should know when and where to do the carving. An important next step in this direction may be to determine if other eating disorders (e.g., anorexia nervosa or binge eating disorder) are in fact also taxonic and categorically distinct from bulimia nervosa. It is possible that the taxon may be for eating disorders in general, meaning that all of these disorders differ qualitatively from normal body and eating concerns, but that the various disorders are not qualitatively different from one another (i.e., they occur on a continuum).

A second reason why the issue of taxonicity versus dimensionality is important concerns how the disorder has been and continues to be studied. If bulimia nervosa is indeed determined to be taxonic, it will be important to consider that persons with bulimia nervosa differ in kind or perhaps in degree and kind from persons with more normative eating and dieting concerns. What this means is that one cannot study the symptomatology of bulimia nervosa without actually studying persons that possess the taxon. Studies of subclinical bulimia or normative forms of dieting, eating problems, or body-weight concerns may be investigating a phenomena that is qualitatively different from (and, hence, not generalizable to) the phenomenon of interest.

Furthermore, if bulimia nervosa is indeed taxonic, it will also be important to consider that sample composition (i.e., purely nonclinical vs. purely clinical vs. mixed) may radically affect the results of research. Several examples of this effect can be seen in recent research. For example, there is a clear body of research linking bulimic behavior with dieting behavior (e.g., Polivy & Herman, 1985; Tuschl, 1990) and body concerns (e.g., Rosen, 1992). That is, bulimic behavior has been found to be correlated with dieting and body concerns. However, the vast majority of these findings have been based primarily on nonclinical samples (in which, as in the current study, the percentage of participants with clinical eating disorders is very low). In two recent factor analytic studies conducted purely within samples of persons diagnosed as having bulimia nervosa (Gleaves & Eberenz, 1995; Gleaves et al., 1993), bulimic behavior was not found to be highly correlated with either dieting behaviors or body concerns. Although these findings were puzzling to some researchers, they are what one would expect if bulimia were actually taxonic. This effect is especially clear when one also considers that Varnado, Williamson, and Netemeyer (1995) used a similar factor analytic methodology and found bulimic behavior to be moderately to highly correlated with both restrictive eating (r = .56) and body image (r = .24) among a sample of college women. The authors noted that, of their 200 participants, 13 met the diagnostic criteria for an eating disorder. Thus, this study used a mixed sample, and the results were what would be expected if bulimia was taxonic.

9 This conclusion is based on the assumption that 13 outliers in the sample could dramatically increase correlations among indicators that were otherwise essentially uncorrelated. That such an inflation can occur can be easily demonstrated. See Paulhus, Bruce, and Trapnell (1995) for an example of this effect.
severity in a nonclinical sample. The point is that we may frequently find inflated correlations between various features of the disorder by examining mixed samples.

A final reason concerns the etiology or causality of the disorder. As noted above, identification of a taxonomic structure for a condition does not necessarily imply a genetic basis. However, the identification of a taxonomic structure may at least give a genetic etiology a higher antecedent probability than if the disorder was determined to be simply at an extreme point on one or more continuums (Meehl, 1992). Thus, as discussed by Meehl in the context of schizotypy, continuing a search for a genetic basis of bulimia nervosa may be more rationally supported if it is determined to be taxonic than if it was found to be clearly dimensional. There may also be other ways in which a taxonic structure is related to etiology other than through genetics. The concept of the environmental mold needs further investigation as does the possibility that the taxon may represent different types of specific environmental events (e.g., the type of pathogenic mothering described by Bruch, 1973) or even types of dieting behavior. For instance, although bulimics do not differ that much from nonbulimic current dieters in the intensity of their current dieting behavior (Lowe et al., 1996), most bulimics lose far more weight in the process of developing their disorder (Garner & Fairburn, 1988) than restrained, non-eating-disordered college students ever do (Lowe, 1984).

In conclusion, this study was the first to apply taxometric procedures to the question of the continuity and discontinuity models of bulimia nervosa. The majority of the results were suggestive of a latent taxon, which is most consistent with the discontinuity model. Future research should try to replicate the finding with other taxometric methods and assessment instruments or strategies of assessment, or both (e.g., interviews rather than self-report). If continued evidence of taxonicity is found, the next step is to learn more about those who possess the taxon, with the goals of better assessment, treatment, and prevention of the disorder.

References


