Diagnosing Major Depressive Disorder XI

A Taxometric Investigation of the Structure Underlying DSM-IV Symptoms

John Ruscio, PhD, Mark Zimmerman, MD,† Joseph B. McGlinchey, PhD,† Iwona Chelminski, PhD,† and Diane Young, PhD†*

Abstract: Psychopathologists have long debated the latent structure of mental disorders, and a number of researchers have suggested that depression may be best characterized as a continuous, rather than categorical, phenomenon. Nonetheless, attention has been drawn to limitations permeating existing research and the need for studies using more appropriate statistical methods developed expressly to tease apart taxonic (categorical) and dimensional (continuous) structural models. The present study examined the structure underlying the DSM-IV symptoms of major depressive disorder in a large outpatient sample rigorously assessed using semistructured clinical interviews. The results of a series of taxometric procedures and consistency tests supported a taxonic structural model, consistent with the only previous taxometric study of DSM-IV symptoms in an adult outpatient sample. In addition to the need for further replication and clarification, these results have implications for the assessment and diagnosis of major depressive disorder. Suggestions for several additional avenues of research are discussed.

Key Words: Major depressive disorder, depression, mood disorder, taxometrics, latent structure.

(*J Nerv Ment Dis* 2007;195: 10–19)

One of the most controversial of nosological disputes involves the latent structure of mental disorders (Grove and Andreasen, 1989; Klein and Riso, 1993; Meehl, 1992). Traditionally, mental disorders have been conceptualized as disease entities that represent a qualitative distinction from normal functioning (e.g., Goodwin and Guze, 1989; Guze, 1992; Robins and Helzer, 1986). On the other hand, many researchers have argued that some, and perhaps many, mental disorders exist along continua with normality (e.g., Mirowsky, 1994; Widiger, 1997). The debate between proponents of categorical (taxonic) and continuous (dimensional) structural models is contentious because it raises theoretical and practical questions rooted in the fundamental nature of psychopathology. Should a disorder be diagnosed as present or absent, or should diagnoses take the form of continuous ratings along one or more dimensions? Because latent structure has implications for the assessment, study, and treatment of mental disorders, it represents one of the most pressing empirical questions facing the developers of the next generation of diagnostic manuals such as the DSM and ICD (Widiger and Clark, 2000).

Flett et al. (1997) described four approaches that have been used to garner evidence relevant to the continuity controversy in the area of depression—phenomenological, typological, etiological, and psychometric—and concluded that research generally supported a dimensional model of depression. However, Flett et al. (1997) also drew attention to some serious methodological limitations of many of the studies that they reviewed. In particular, they argued that investigators often used inappropriate statistical methods to test latent structure and called for research using techniques such as Meehl's (1995) taxometric method, which was developed expressly for this purpose. In response to this suggestion, a number of investigators have performed taxometric investigations of depressive constructs.

The philosophy of science motivating the taxometric method and a number of its attractive features have been discussed in many sources (e.g., Meehl, 1995; Meehl and Golden, 1982; Waller and Meehl, 1998), including a recent special section of the *Journal of Abnormal Psychology* (Cole, 2004), and taxometric tutorials have appeared as well (e.g., Ruscio, In press; Ruscio and Ruscio, 2004b, 2004c). Ruscio and Ruscio (2004a) discuss the special role that taxometrics can play in research on the boundary conditions of psychopathology. Whereas some taxometric studies have explored the structure of putative subtypes or variants of major depressive disorder (MDD; Ambrosini et al., 2001; Beach and Amir, 2003; Grove et al., 1987; Haslam and Beck, 1994), others have targeted MDD itself. Hankin et al. (2005) obtained dimensional results in an adolescent sample, as did Ruscio and Ruscio (2000) in two adult samples, one clinical and one more broadly medical. However, a more recent study by Ruscio et al. (2005a) suggested that at least some sources

*Psychology Department, The College of New Jersey, Ewing New Jersey, and †Department of Psychiatry and Human Behavior, Brown University School of Medicine, Rhode Island Hospital, Providence, Rhode Island.

Send reprint requests to John Ruscio, PhD, Psychology Department, The College of New Jersey, PO Box 7718, Ewing, NJ 08628.

ISSN: 0022-3018/07/19501-0010

DOI: 10.1097/01.nmd.0000252025.12014.c4

of self-report depression data can yield less interpretable taxometric results than the ratings of clinical interviewers. Moreover, the latter data source produced evidence of taxonic structure. Because Ruscio and Ruscio (2000) relied exclusively on self-report data, their dimensional findings may reflect a failure to detect an MDD taxon due to limitations of the data available for their analyses.

In addition to studying the differential utility of alternate data sources, Ruscio et al. (2005a) constructed composite indicators for analysis based more closely on DSM-IV criteria (American Psychiatric Association, 1994) for MDD than has been the case in previous taxometric research, worked with a large outpatient sample, and applied a number of interpretive aids. Nonetheless, the evidence of taxonic structure obtained in this single investigation cannot resolve such an important question, and the latent structure of MDD requires additional study. The present study was designed as a conceptual replication of Ruscio et al. (2005a). Clinical interview data from another large sample of outpatients were used to construct indicators based on DSM-IV diagnostic symptoms for MDD and perform a series of taxometric analyses. The central research question tested here is whether MDD should be diagnosed as a discrete mental disorder or indexed using continuous ratings along one or more dimensions.

METHODS

Data Source

Data for the present study were collected as part of the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project (Zimmerman, 2003), which integrates research methodology into a community-based outpatient practice affiliated with an academic medical center. Details regarding referral, recruitment, and other issues related to the MIDAS project have been provided elsewhere (e.g., Zimmerman and Mattia, 1999; Zimmerman et al., 2005a). The present sample includes all 1800 patients who have been evaluated with a semistructured diagnostic interview in the Rhode Island Hospital Department of Psychiatry outpatient practice. The majority of the subjects were white (87.3%, $N = 1571$), female (61.4%, $N = 1106$), and married (40.8%, $N = 735$) or single (31.0%, $N = 558$). The most frequent current DSM-IV diagnoses were MDD (46.0%, $N = 829$), social phobia (28.8%, $N = 519$), panic disorder (18.4%, $N = 331$), and generalized anxiety disorder (17.8%, $N = 320$).

All patients were interviewed by a diagnostic rater who administered the Structured Clinical Interview for DSM-IV (SCID; First et al., 1995) and the Schedule for Affective Disorders and Schizophrenia (SADS; Endicott and Spitzer, 1978). Detailed procedural and reliability information are presented elsewhere (Zimmerman et al., 2005a). Composite variables representing each of the DSM-IV criteria for major depression were constructed from interviewer ratings, including both trichotomous (present, subthreshold, absent) SCID ratings and scaled (0–5, with verbal descriptions of each) SADS ratings. Table 1 lists the variables included in each composite as well as its internal consistency.

The taxometric method requires that variables used as indicators of the target construct must distinguish the putative groups (i.e., MDD and non-MDD patients) with sufficient validity and be substantially less correlated within groups than in the full sample (Meehl, 1995; Ruscio et al., 2006). Rather than automatically including each DSM-IV symptom of MDD as an indicator, in which case ambiguous results may emerge, they were screened to ensure that both data requirements were met. To address the first requirement, validity was estimated by examining the separation of MDD+ and MDD– patients on each composite variable. Variables representing sleep and psychomotor disturbance, as well as suicidality, poorly distinguished these diagnostic groups (Cohen $d < 1.00$). This left a pool of seven variables that appeared sufficiently valid to warrant further consideration. To address the second requirement, these variables' correlations were examined in the full sample as well as within MDD+ and MDD– groups. A distinct cluster of variables—depressed mood, weight change, worthlessness/guilt, and hopelessness/helplessness—was correlated nearly as highly within groups as in the full sample. Although not conceptually redundant, this empirical redundancy could pose problems for taxometric analyses. Therefore, these four variables were collapsed into a single indicator. This yielded a final set of four valid, nonredundant indicators for taxometric analysis: the collapsed variable described above, loss of interest, fatigue, and impaired concentration.

The indicators included six of the nine DSM-IV symptom criteria for MDD as well as two associated features. This provides some assurance that, despite the removal of three variables and the combination of others to meet the data requirements for taxometric analysis, the indictor set submitted to analysis does represent the MDD construct outlined in DSM-IV. Complete data were available on these indicators for 1748 patients. The removal of 10 patients whose depression was better accounted for by a general medical condition resulted in a final sample size of 1738. Summary statistics and validity estimates for the four indicators, each of which was standardized prior to analysis, are shown in Table 2.

Data Analysis

The taxometric method is grounded in the examination of the consistency of results across multiple analytic techniques. Because expert opinion has not coalesced regarding the taxometric results that best differentiate taxonic and dimensional structural models, and there are scant data bearing on the incremental validity of structural inferences reached using additional consistency tests, a broad approach was adopted. Data were submitted to four taxometric procedures: MAMBAC (Meehl and Yonce, 1994), MAXCOV (Meehl and Yonce, 1996), MAXEIG (Waller and Meehl, 1998), and L-Mode (Waller and Meehl, 1998). The rationale, mathematics, and implementation options for each procedure are provided in the primary sources listed as well as taxometric tutorials cited. The specific implementation of each procedure is described in the Appendix.

To complement these procedures—each of which yielded one or more curves whose shapes were inspected for evidence of taxonic or dimensional structure—both traditional and newer quantitative indices were calculated. The *SD* of taxon base rate estimates has been proposed as one way to

TABLE 1. Composites Representing DSM-IV Symptoms (and Associated Features) of Major Depression

Symptom (# items)	Items	Standardized Item α
Depressed mood (2)	Depressed mood (T)	.91
	Subjective feelings of depression (S)	
Diminished interest (4)	Diminished interest, all or most activities (T)	.93
	Diminished interest, several activities (T)	
	Pervasiveness of diminished interest (S)	
	Diminished drive to perform (T)	
Appetite/weight change (8)	Decreased appetite (T, S)	.62
	Increased appetite (T, S)	
	Weight loss (T, S)	
	Weight gain (T, S)	
Sleep disturbance (4)	Insomnia, initial (T)	.82[a]
	Insomnia, middle (T)	
	Insomnia, terminal (T)	
	Insomnia (S)	
Psychomotor disturbance (4)	Psychomotor agitation (T, S)	.97[b]
	Psychomotor retardation (T, S)	
Fatigue (2)	Fatigue (T, S)	.96
Worthlessness/guilt (4)	Worthlessness (T, S)	.88
	Guilt (T, S)	
Impaired concentration (4)	Diminished ability to concentrate (T, S)	.88
	Indecisiveness (T, S)	
Suicidality (5)	Thoughts of death (T)	.84
	Suicidal ideas (T)	
	Suicide plan (T)	
	Suicide attempt (T)	
	Suicidal tendencies (S)	
Hopelessness/helplessness (3)[c]	Hopelessness (T, S)	.80
	Helplessness (T)	

T, Trichotomous item (present, subthreshold, absent); S, scale (0–5 rating).
[a]When two additional items assessing hypersomnia (one T, one S) were included, $\alpha = .57$.
[b]Whereas $\alpha = .69$ for the four-item composite, $\alpha = .97$ for both an agitation and a retardation composite.
[c]Hopelessness and helplessness are features often associated with major depression, not DSM-IV symptoms.

TABLE 2. Summary Statistics and Validity Estimates for the Four Indicators[a]

	Depressed Mood, Weight Change, Worthlessness/ Guilt, Hopelessness/ Helplessness	Loss of Interest	Fatigue	Impaired Concentration
Skew	–.13	–.40	–.75	.00
Kurtosis	–.93	–1.49	–1.01	–1.37
MDD+ *M*	.70	.70	.53	.57
MDD+ *SD*	.63	.56	.64	.79
MDD– *M*	–.62	–.62	–.47	–.50
MDD– *SD*	.84	.89	1.02	.89
Validity (*d*)	1.76	1.75	1.16	1.27

[a]$N = 1738$; MDD+ $N = 812$; MDD– $N = 926$. All indicators have been standardized in the full sample. Indicator validity is expressed as Cohen *d*, the mean difference between groups standardized using pooled within-groups variances (weighted by *df*).

assess consistency of results with a taxonic model (i.e., smaller *SD*s hint at the existence of a taxon; Waller and Meehl, 1998), and these values are presented even though a recent Monte Carlo study suggests that they may not distinguish taxonic and dimensional structure well (Ruscio, In press). Likewise, the goodness of fit index (GFI; Waller and Meehl, 1998) is presented even though questions have been raised about the utility of this index in differentiating structural models and the appropriateness of a fixed threshold value of ≥.90 as indicative of taxonic structure (Cleland et al., 2000; Haslam and Cleland, 2002; Ruscio et al., In press).

In addition, relatively recently developed interpretive aids were used to supplement—not to replace—the subjective evaluation of curve shapes. Because a number of factors can influence the shape of taxometric curves, Ruscio et al. (2004) proposed that researchers perform parallel analyses of comparison data simulated to reproduce the distributional and correlational characteristics of the research data using both taxonic and dimensional structural models. This yields sampling distributions of results for known latent structures, to which the results for the research data can be compared.

The data simulation procedure used here was outlined in Ruscio et al. (2004), described in detail in Ruscio et al. (2006), illustrated in Ruscio and Ruscio (2004b, 2004c), and studied extensively in Ruscio et al. (In press). Details on the

simulation of comparison data for the present study are provided in the Appendix. Evidence suggests that the use of simulated comparison data can be a helpful supplementary tool in taxometric investigations. In particular, they afford the calculation of a comparison curve fit index (CCFI) to quantify the match between the results for research data and those of taxonic and dimensional comparison data (Ruscio et al., In press; see Appendix for details). In many cases, a simple visual inspection of a three-panel figure (results for research data, simulated taxonic data, and simulated dimensional data) reveals clear support for one or the other structural model under consideration. In other cases, the CCFI can shed some light on more subtle features of the graphs. A preliminary study found that precursors to the CCFI yielded structural inferences as accurate as visual inspection (Ruscio, 2004). More recent and large-scale Monte Carlo research, using the updated version of the simulation programs employed here, found that interpretations based on the CCFI were much more accurate than those based on more traditional indices (e.g., *SD* of taxon base rate estimates, GFI; Ruscio et al., In press).

RESULTS

Taxometric Analyses

In the MAMBAC analyses, nine of the 12 curves contained peaked shapes consistent with taxonic structure, two contained ambiguously peaked curves (i.e., the apparent peak was near the end of the graph), and the remaining curve was ambiguously jagged. The averaged curve, shown in Figure 1, was considerably more similar to that for taxonic than dimensional comparison data. The large CCFI, small *SD* of base rate estimates, and high GFI (Table 3), provided further evidence of taxonic structure.

MAXCOV analyses yielded less conclusive results. Five of the 12 curves were peaked (consistent with taxonic structure), five were fairly flat (consistent with dimensional structure), and the remaining two curves were ambiguous. The averaged curve was a bit more similar to that for taxonic comparison data, but not as obviously as was the case for the MAMBAC results. The CCFI, *SD* of base rate estimates, and GFI also supported an inference of taxonic structure.

MAXEIG results were even more ambiguous. One of the four curves was strongly peaked (consistent with taxonic structure) and one was flat (consistent with dimensional structure). The remaining two curves contained left-side plateaus that declined toward the right. These final two curves may represent the untoward influence of nontrivial indicator correlations within the putative non-MDD group. Because MAXEIG includes all indicators in each analysis, it may be more sensitive to within-group correlations. Such an interpretation is speculative at best, however, and this panel of curves did not provide much useful evidence bearing on a structural conclusion. As with MAXCOV, the results for comparison data did not shed much light on MAXEIG results. Whereas the CCFI weakly supported taxonic structure for MAXCOV, it was ambiguous for MAXEIG. The *SD* of base rate estimates was quite low, and the GFI above the recommended threshold, so to the extent that one chooses to rely on these indices, they provide support for taxonic structure.

L-Mode results arguably provide some further evidence of taxonicity. Although the distribution of factor scores did not contain a central trough, it appears that two groups' distributions may have been mixed together. A distinct mode toward the right side tapers toward a more rounded, and less distinct, mode toward the left. As with the MAXEIG curves described, this is a somewhat speculative interpretation. The GFI was at its highest for the L-Mode analysis.

Taken together, these analyses provide some evidence that is highly consistent with taxonic structure (all MAMBAC results), some that is more moderately consistent with taxonic structure (portions of the results for the remaining three procedures), and comparatively little evidence consistent with dimensional structure. Because the inclusion of 103 patients meeting criteria for bipolar disorder may have complicated the interpretation of results—bipolar disorder shares some, but not all, features with MDD—follow-up analyses were performed after excluding these patients. The removal of patients meeting criteria for bipolar disorder had little effect on the results. Although it is unfortunate that some results remained ambiguous, it is reassuring to observe that the inclusion versus exclusion of bipolar patients would not alter conclusions drawn about the latent structure of MDD.

Classification of Patients into Groups

Given that the structural analyses suggested taxonic structure, patients were assigned to taxon (MDD+) and complement (MDD–) groups to examine the diagnostic status and symptom patterns of group members. In the taxometric literature, cases are classified most often using the results of a MAXCOV analysis. Specifically, the taxon base rate and the valid positive (VP) and false positive (FP) rates achieved by the optimal cutting score on each indicator are estimated from the MAXCOV results, and then the Bayes theorem is used to calculate the probability of taxon membership for each case. Cases are assigned to the more probable class. In addition, the results of L-Mode analyses can be used to classify cases in a different way. Each case's score on the latent factor is estimated and then compared with the scores beneath the two modes in the distribution. Cases are assigned to groups based on the smaller of their two absolute distances to these modes (for further details on both classification methods, see Waller and Meehl, 1998). The agreement between the classifications provided by the MAXCOV and L-Mode analyses was fairly high, $\kappa = .82$. This level of classification agreement provides reasonable assurance that the taxometric analyses consistently identified the same taxon.

These classifications include the 1738 cases with complete data on the four indicators submitted to taxometric analysis, and they do not take into account symptoms that were not included in the analyses. To address both of these limitations, an iterative technique was applied to classify all 1800 cases in the sample using the nine DSM-IV symptoms of MDD. To begin, the VP and FP rates for each of the dichotomous symptom-level variables were estimated within groups obtained from a taxometric analysis; the taxon base rate was estimated by the taxometric procedure. The Bayes theorem was then used to calculate a probability of taxon membership for each of the 1800 cases in the sample using

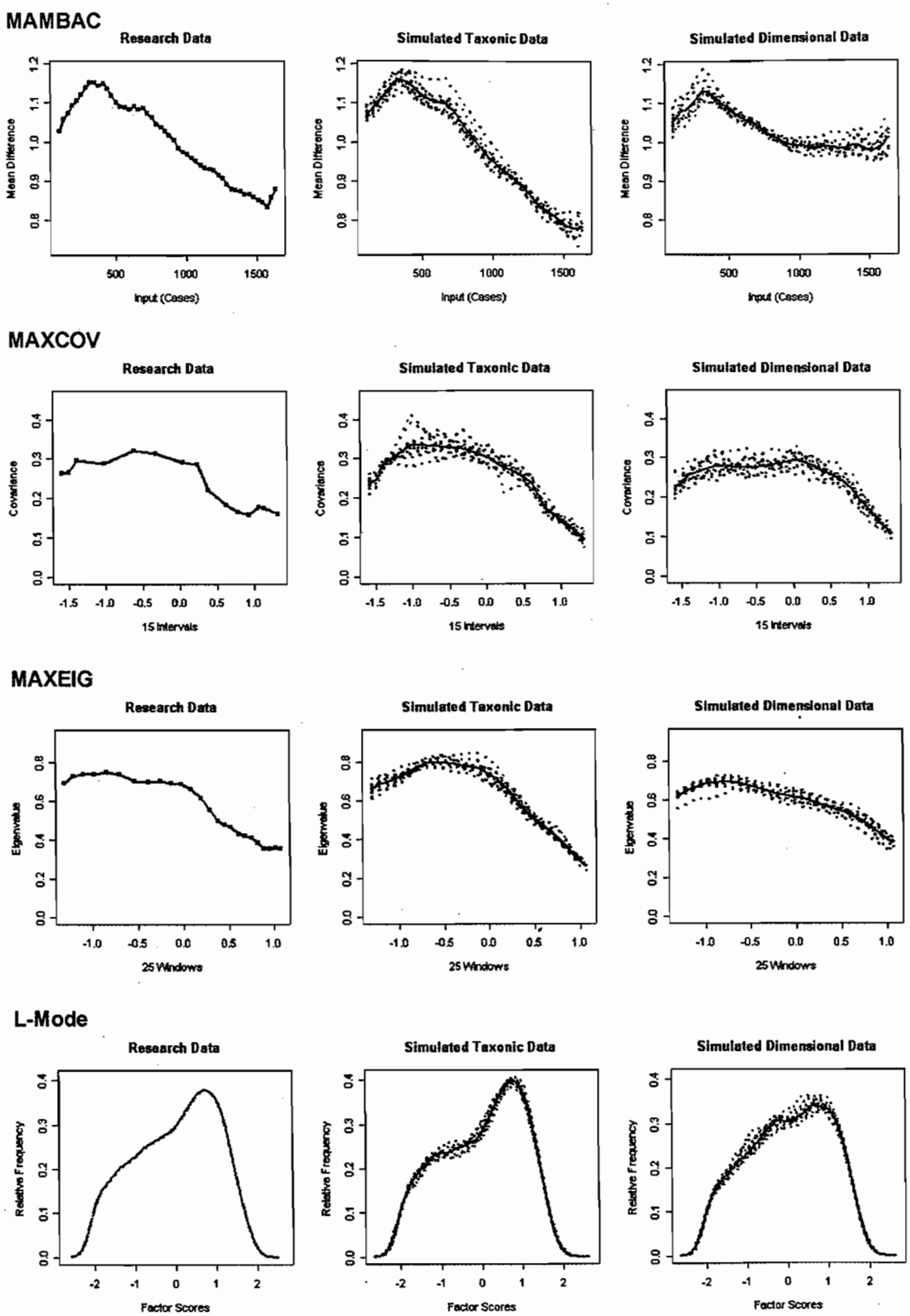

FIGURE 1. Taxometric curves for each procedure. MAMBAC, MAXCOV, and MAXEIG curves are averaged across full panels of results. For simulated comparison data plots, the 10 dotted lines represent curves for individual data sets and solid lines represent their average.

TABLE 3. Summary of Analytic Results

Taxometric Procedure	*M* (*SD*) of Taxon Base Rate Estimates	GFI[a]	CCFI[b]
MAMBAC	.53 (.08)	.95	.66
MAXCOV	.62 (.09)	.95	.55
MAXEIG	.69 (.03)	.91	.50
L-Mode	.62/.61[c]	.96	—

[a]GFI values above .90 are interpreted as evidence supportive of taxonic structure.
[b]CCFI values above .50 are interpreted as evidence supportive of taxonic structure.
[c]L-Mode base rate estimates are (1) the average value of the estimates calculated from the location of each apparent mode and (2) an empirical estimate calculated as the proportion of cases assigned to the taxon.

the taxon base rate estimate and the nine symptom-level variables, and each case was assigned to the more probable group. This procedure was repeated (i.e., the taxon base rate and the VP and FP rates were estimated using the updated classification of cases, and the Bayes theorem was used to assign cases to groups) until all cases were classified into the same groups on two successive iterations. When this technique began with the classification provided by the MAXCOV analysis, seven iterations were required to converge. When this technique began with the classification provided by the L-Mode analysis, convergence was achieved with three iterations. Notably, the final classifications of cases were identical when beginning with MAXCOV and L-Mode classifications: the taxon contained the same 1057 (58.7%) of the 1800 cases. In other words, MAXCOV and L-Mode yielded reasonably similar classifications based on the four indicators submitted to taxometric analysis (as noted, $\kappa = .82$), but iterative updating using the nine DSM-IV symptoms yielded results that agreed perfectly. This final classification of cases is used below.

Table 4 shows the diagnoses assigned to taxon and complement members. Virtually all taxon members (99.7%) met criteria for a mood disorder, whereas relatively few complement members did so (20.9%). Taxon members most often met criteria for MDD (78.3%), with 8.2% or less meeting criteria for each of the other mood disorders. Complement members most often met criteria for the catch-all category of depressive disorder not otherwise specified (11.7%), which included roughly half of all complement members with a diagnosable mood disorder; patients in this category constituted a much smaller proportion of the taxon. Most patients diagnosable with bipolar disorder were classified into the taxon, and patients diagnosable with dysthymia were somewhat more likely to be classified into the taxon. Too few patients were diagnosable with cyclothymia to afford a comparison. These diagnostic data suggest that the taxon uncovered in the taxometric analyses consisted largely, although not exclusively, of patients meeting criteria for MDD, whereas the complement group contained a single patient meeting criteria for MDD but a number of individuals with other mood disorders. Table 5 shows this by crossing taxon membership with MDD diagnostic status ($\kappa = .75$).

In addition to these diagnostic comparisons, group differences on variables relevant to the DSM-IV symptom criteria for MDD were examined. Table 6 shows the percentage of taxon and complement members rated above threshold ("present" as opposed to "absent" or "subthreshold") on each symptom or associated variable. Not surprisingly, the findings are highly similar to those of the more extensive psychometric evaluation by Zimmerman et al. (2005a) of DSM-IV symptom criteria for diagnosing MDD in the MIDAS sample. This lends further support to the notion that the taxon uncovered by the taxometric analyses corresponds to MDD. The percentages shown for the nine DSM-IV symptoms (bold print) represent the VP and FP rates used in the final classification of cases to groups. Provided that one has an estimate of the taxon base rate in a particular sample, one could use these VP and FP estimates to apply the Bayes theorem in the assignment of cases to taxon and complement as an alternative to the DSM-IV's diagnostic algorithm.

TABLE 4. Number of Taxon and Complement Members Meeting DSM-IV Mood Disorder Criteria

Disorder	Taxon (*N* = 1057)	Complement (*N* = 743)
Any mood disorder	1054 (99.7%)	155 (20.9%)
Major depressive disorder	828 (78.3%)	1 (0.1%)
Bipolar disorder	87 (8.2%)	18 (2.4%)
Dysthymia	80 (7.6%)	47 (6.3%)
Cyclothymia	1 (0.1%)	2 (0.3%)
Depressive disorder, not otherwise specified	58 (5.5%)	87 (11.7%)

DISCUSSION

In the present analysis of the MIDAS sample, taxometric curve shapes, traditional consistency tests such as the *SD* of taxon base rate estimates and the GFI, and the interpretive aids provided by supplementary analyses of taxonic and dimensional comparison data provided evidence in support of the taxonic structure of MDD. This is consistent with the findings of Ruscio et al. (2005a), the only other taxometric study of MDD in a large outpatient sample. This replication provides support for the existence of a boundary between patients diagnosable with MDD and patients who are not experiencing a major depressive episode. However, it is also the case that neither of these two studies yielded unambiguously taxonic results. In the present study, the results of one taxometric procedure (MAMBAC) were highly consistent with expectations for taxonic structure, with those for three other procedures generally—but not as strongly—supportive of the same conclusion. Thus, further attempts to replicate are warranted.

TABLE 5. Cross-Classification Table for Taxon Membership and MDD Diagnostic Status

	Diagnostic Status	
	MDD+	MDD–
Taxon	828	229
Complement	1	742

TABLE 6. Percentage of Taxon and Complement Members Meeting DSM-IV Diagnostic Criteria[a]

Variable	Taxon (N = 1057)	Complement (N = 743)
Depressed mood	**88.8**	**8.6**
Diminished interest	**76.5**	**5.3**
Appetite/weight change	**67.0**	**19.2**
ased appetite	41.6	9.5
Increased appetite	19.2	6.8
Weight loss	20.7	5.8
Weight gain	16.1	3.8
Sleep disturbance	**82.0**	**29.0**
Insomnia, initial	45.4	12.3
Insomnia, middle	51.6	16.5
Insomnia, terminal	28.7	10.0
Hypersomnia	18.7	5.4
Psychomotor disturbance	**53.7**	**9.6**
Psychomotor agitation	34.9	7.9
Psychomotor retardation	26.6	1.9
Fatigue	**87.6**	**24.7**
Worthlessness/guilt	**72.7**	**16.6**
Worthlessness	58.4	7.9
Guilt	52.1	12.1
Impaired concentration	**85.3**	**22.7**
Diminished concentration	80.4	20.7
Indecisiveness	50.0	6.2
Suicidality	**54.1**	**11.5**
Thoughts of death	53.4	10.0
Suicidal ideas	28.8	3.4
Suicide plan	14.9	1.1
Suicide attempt	3.0	1.6

[a]DSM-IV symptoms appear in bold.

Despite the convergence on taxonic structure in these two studies, the boundary separating members of an MDD taxon from nonmembers is in need of further clarification. Whereas Ruscio et al. (2005a) found that the estimated taxon base rate matched the diagnosed rate of MDD in their sample, the taxon uncovered in the present study was broader. Along with nearly all patients diagnosed with MDD, the taxon included a nontrivial number of patients with other mood disorders. One possible explanation for this apparent discrepancy is that the indicators submitted to analysis in the present study were more highly correlated within the putative complement (non-MDD patients) than the taxon. This would be expected to increase taxon base rate estimates obtained from taxometric procedures through a predictable influence on curve shapes (i.e., raising the left side of each curve; Meehl and Yonce, 1994, 1996). Indicator correlations were more comparable across groups in Ruscio et al. (2005a), and therefore their estimates of the taxon base rate—and the classifications of cases that depend in part on these estimates—may be more reliable. Future research is needed to determine whether an MDD taxon consists of a proportion of a sample that is similar to or larger than the proportion meeting DSM-IV criteria for MDD.

In addition to clarifying the taxon's boundary conditions, several other lines of research appear promising. First, psychometric evaluations of potentially diagnostic signs and symptoms can be performed after assigning cases to taxon (MDD+) and complement (MDD–) groups. The DSM-IV diagnostic algorithm may not be maximally efficient in identifying MDD, and the developers of future editions of diagnostic manuals may wish to revise the algorithm. The existence of a taxonic boundary suggests that psychometric evaluations of diagnostic criteria should begin by assigning cases to groups that most closely correspond to taxon and complement membership; adhering to the DSM-IV diagnostic algorithm may yield groups that correspond more poorly to such latent classes. Although the present results suggest that some of the DSM-IV symptom criteria more validly distinguish taxon and complement members than do others, this does not necessarily mean that the latter should be excluded from the diagnostic algorithm. Obtaining interpretable taxometric results requires the use of highly valid indicators, whereas criteria of more modest validity can still make important contributions to diagnosis. Zimmerman et al. (2005b) discuss strategies for determining which criteria should be included and excluded from a diagnostic algorithm. In the present study, a technique was illustrated by which one can test latent structure using a set of indicators chosen to meet the requirements of taxometric analysis and then bridge the results to a classification of cases that takes advantage of a broader array of signs or symptoms.

Second, on a related note, the way in which depression is assessed may benefit from knowledge of its latent structure. Diagnostic algorithms would be most effective if based on the assessment modalities and instruments that most effectively differentiate between taxon and complement members. Assessments that aim to locate individuals' scores along one or more continuous dimensions most accurately do not necessarily afford the most accurate classifications of cases to groups (Meehl, 1992; Ruscio and Ruscio, 2002).

Third, the search for depressive subtypes may be more effective if grounded in a preliminary assignment of cases to the MDD taxon and its complement. Whereas structural studies of MDD require a mixture of MDD+ and MDD– cases, studies of putative subtypes may be more powerful if performed within a sample consisting exclusively of MDD+ cases. A diagnostic algorithm that accurately identifies MDD taxon members could be used to select an appropriate sample for analyses targeting boundaries between proposed subtypes of depression.

Fourth, the developmental trajectory of MDD merits further study. Factors that influence the transition into and out of an MDD taxon should be explored, as well as potential structural differences across age groups. A taxometric study by Hankin et al. (2005), for example, found that a dimensional model provided better fit than a taxonic model in a study of child and adolescent depression. Beauchaine (2003) discusses the role that taxometric studies can play in understanding the developmental course of psychopathology, including the possibility that the most useful structural model varies with age.

Fifth, knowledge of the latent structure of MDD—including potential subtypes—can inform etiological theorizing and research. Assigning patients to appropriate groups can increase the statistical power of etiological research (Fraley and Waller, 1998). Moreover, theories purporting to explain the development of MDD must satisfactorily account for its structure as well. Different causal models are compatible with taxonic and dimensional structures (Haslam, 1997).

Finally, perhaps the most important way to build upon structural investigations is to test for treatment implications. Psychometric, taxometric, or other statistical analyses of diagnoses and classifications may be less likely to influence the refinement of diagnostic manuals and the clinical judgment of practitioners than would data showing differential treatment responses of taxon members. Of course, finding that a taxonic model best represents the latent structure of MDD provides no guarantee that there are any treatment implications, but the exploration of interactions between psychotherapeutic, pharmacological, or combined treatment regimens with diagnostic status merits serious consideration. To the extent that taxon membership corresponds to a meaningful latent discontinuity in depressive phenomena, the search for differential treatment responses may be more fruitful if one uses a diagnostic algorithm that yields taxon membership rather than DSM-IV diagnostic status.

CONCLUSION

Taxometric analyses performed in the MIDAS data set, which consists of a large outpatient sample rigorously assessed using a semistructured clinical interview, suggest that a taxonic model best captures the latent structure of MDD. Using indicator variables based on DSM-IV criteria for MDD, the MAMBAC procedure provided the strongest evidence in favor of taxonic structure. However, the balance of evidence from MAXCOV, MAXEIG, and L-Mode analyses—as well as a number of traditional and newer consistency tests—served to corroborate this conclusion. This finding is consistent with the results of the only other taxometric study of MDD in a large outpatient sample (Ruscio et al., 2005), which also led to the conclusion that the structure underlying DSM-IV symptoms is taxonic. The diagnostic status and symptom patterns of taxon and complement members draw attention to some important differences between taxon membership and DSM-IV diagnostic status. Although these findings warrant additional replication and suggest research avenues for further clarification, they also have implications for the assessment and diagnosis of MDD as well as research into its etiology and treatment.

REFERENCES

Ambrosini P, Bennett DS, Cleland CM, Haslam N (2002) Taxonicity of adolescent melancholia: A categorical or dimensional construct? *J Psychiatr Res.* 36:247–256.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th ed). Washington (DC): American Psychiatric Association.

Beach SRH, Amir N (2003) Is depression taxonic, dimensional, or both? *J Abnorm Psychol.* 112:228–236.

Beauchaine TP (2003) Taxometrics and developmental psychopathology. *Dev Psychopathol.* 15:501–527.

Cleland CM, Rothschild L, Haslam N (2000) Detecting latent taxa: Monte Carlo comparison of taxometric, mixture model and clustering procedures. *Psychol Rep.* 87:37–47.

Cole DA (2004) Taxometrics in psychopathology research: An introduction to some of the procedures and related methodological issues. *J Abnorm Psychol.* 113:3–9.

Endicott J, Spitzer R (1978) A diagnostic interview: The schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry.* 35:837–844.

First MB, Spitzer RL, Gibbon M, Williams JBW (1995) *Structured Clinical Interview for DSM-IV Axis I Disorders—Patient Edition (SCID-I/P, Version 2.0).* New York: Biometrics Research Department, New York State Psychiatric Institute.

Flett GL, Vrendenburg K, Krames L (1997) The continuity of depression in clinical and nonclinical samples. *Psychol Bull.* 121:395–416.

Fraley RC, Waller NG (1998) Adult attachment patterns: A test of the typological model. In JA Simpson, WS Rholes (Eds), *Attachment Theory and Close Relationships* (pp 77–114). New York: Guilford.

Goodwin DW, Guze SB (1989) *Psychiatric Diagnosis* (4th ed). New York: Oxford University Press.

Grove WM, Andreason NC (1989) Quantitative and qualitative distinctions between psychiatric disorders. In LN Robins, JE Barrett (Eds), *The Validity of Psychiatric Diagnoses* (pp 127–139). New York: Raven Press.

Grove WM, Andreasen NC, Young M, Endicott J, Keller MB, Hirschfeld RMA, Reich T (1987) Isolation and characterization of a nuclear depressive syndrome. *Psychol Med.* 17:471–484.

Guze SB (1992) *Why Psychiatry Is a Branch of Medicine.* New York: Oxford University Press.

Hankin BL, Fraley RC, Lahey BB, Waldman ID (2005) Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *J Abnorm Psychol.* 114:96–110.

Haslam N (1997) Evidence that male sexual orientation is a matter of degree. *J Pers Soc Psychol.* 73:862–870.

Haslam N, Beck AT (1994) Subtyping major depression: A taxometric analysis. *J Abnorm Psychol.* 103:686–692.

Haslam N, Cleland C (2002) Taxometric analysis of fuzzy categories: A Monte Carlo study. *Psychol Rep.* 90:401–404.

Klein DN, Riso LP (1993) Psychiatric disorders: Problems of boundaries and comorbidity. In CG Costello (Ed), *Basic Issues in Psychopathology* (pp 19–66). New York: Guilford.

Meehl PE (1992) Factors and taxa, traits and types, differences of degree and differences in kind. *J Pers.* 60:117–174.

Meehl PE (1995) Bootstraps taxometrics: Solving the classification problem in psychopathology. *Am Psychol.* 50:266–275.

Meehl PE, Golden RR (1982) Taxometric methods. In PC Kendall, JN Butcher (Eds), *Handbook of Research Methods in Clinical Psychology* (pp 127–181). New York: Wiley.

Meehl PE, Yonce LJ (1994) Taxometric analysis, I: Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychol Rep.* 74:1059–1274.

Meehl PE, Yonce LJ (1996) Taxometric analysis, II: Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV procedure). *Psychol Rep.* 78:1091–1227.

Mirowsky J (1994) The advantages of indexes over diagnoses in scientific assessment. In WR Avison, IH Gotlib (Eds), *Stress and Mental Health: Contemporary Issues and Prospects for the Future* (pp 239–258). New York: Plenum.

Robins LN, Helzer JE (1986) Diagnosis and clinical assessment: The current state of psychiatric diagnosis. *Ann Rev Psychol.* 37:409–432.

Ruscio J (2004) *Bootstrapping Sampling Distributions to Quantify the Fit of Taxometric Curves.* Paper presented at the annual meeting of the American Psychological Society, Chicago (IL).

Ruscio J (In press) Taxometric analysis: An empirically grounded approach to implementing the method. *Criminal Justice and Behavior.*

Ruscio J, Brown TA, Ruscio AM (2005) A taxometric investigation of DSM-IV major depression in a large outpatient sample: Interpretable structural results depend on the mode of assessment. Manuscript submitted for publication.

Ruscio J, Haslam N, Ruscio AM (2006) *Introduction to the Taxometric Method: A Practical Guide.* Mahwah (NJ): Lawrence Erlbaum Associates.

Ruscio J, Ruscio AM (2000) Informing the continuity controversy: A taxometric analysis of depression. *J Abnorm Psychol.* 109:473–487.
Ruscio J, Ruscio AM (2002) A structure-based approach to psychological assessment: Matching measurement models to latent structure. *Assessment.* 9:4–16.
Ruscio J, Ruscio AM (2004a) Clarifying boundary issues in psychopathology: The role of taxometrics in a comprehensive program of structural research. *J Abnorm Psychol.* 113:24–38.
Ruscio J, Ruscio AM (2004b) A conceptual and methodological checklist for conducting a taxometric investigation. *Behav Ther.* 35:403–447.
Ruscio J, Ruscio AM (2004c) A nontechnical introduction to the taxometric method. *Understand Stat.* 3:151–194.
Ruscio J, Ruscio AM, Keane TM (2004) Using taxometric analysis to distinguish a small latent taxon from a latent dimension with positively skewed indicators: The case of Involuntary Defeat Syndrome. *J Abnorm Psychol.* 113:145–154.
Ruscio J, Ruscio AM, Meron M (In press) Applying the bootstrap to taxometric analysis: Generating empirical sampling distributions to help interpret results.
Waller NG, Meehl PE (1998) *Multivariate Taxometric Procedures: Distinguishing Types From Continua.* Thousand Oaks (CA): Sage.
Widiger TA (1997) Mental disorders as discrete clinical conditions: Dimensional versus categorical classification. In SM Turner, M Hersen (Eds), *Adult Psychopathology and Diagnosis* (pp 3–23). New York: Wiley.
Widiger TA, Clark LA (2000) Toward DSM-V and the classification of psychopathology. *Psychol Bull.* 126:946–963.
Zimmerman M (2003) Integrating the assessment methods of researchers in routine clinical practice: The Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project. In M First (Ed), *Standardized Evaluation in Clinical Practice* (Vol 22, pp 29–74). Washington (DC): American Psychiatric Publishing, Inc.
Zimmerman M, Mattia JI (1999) Psychiatric diagnosis in clinical practice: Is comorbidity being missed? *Compr Psychiatry.* 40:182–191.
Zimmerman M, McGlinchey J, Young D, Chelminski I (2005a) Diagnosing major depressive disorder, I: A psychometric evaluation of the DSM-IV symptom criteria. *J Nerv Ment Dis.* 194:158–163.
Zimmerman M, McGlinchey J, Young D, Chelminski I (2005b) Diagnosing major depressive disorder, III: Can some symptoms be eliminated from the diagnostic criteria? *J Nerv Ment Dis.* 194:313–317.

APPENDIX

Implementing the Taxometric Procedures

MAMBAC was performed by placing cuts between each successive case along an input indicator, beginning and ending 100 cases from each extreme, and calculating the mean difference on an output indicator. By using indicators in all input/output indicator configurations, this resulted in 12 MAMBAC curves. Each curve's shape was inspected for its consistency with taxonic or dimensional structure, and each curve was also used to calculate an estimate of the putative taxon's base rate in the sample. Finally, by assigning cases to putative taxon and complement groups using a base-rate classification technique (Ruscio et al., 2006), the GFI was calculated using the formula in Waller and Meehl (1998).

MAXCOV was performed by dividing cases into 15 equal-sized intervals along an input indicator and calculating the covariance of two output indicators. By using indicators in all input/output configurations, this resulted in 12 MAXCOV curves. Once again, curve shapes were inspected and base rate estimates were calculated. Following the procedure provided in Meehl and Yonce (1996), cases were classified into putative taxon and complement groups using the Bayes theorem, and the GFI was calculated.

MAXEIG, a multivariate extension of MAXCOV, was performed by dividing cases into 25 windows that overlapped 90% with their neighbors and calculating the first (largest) eigenvalue of the covariance matrix (a variance-covariance matrix with the variances along the diagonal replaced by zeros) of the remaining three output indicators. This resulted in four MAXEIG curves, each of which was inspected and used to calculate a base rate estimate. As with MAXCOV, cases were classified using the Bayes theorem and the GFI was calculated.

L-Mode involves a factor analysis of all available indicators, with the distribution of scores on the first factor calculated. The shape of the distribution of these scores was inspected and used to estimate the taxon base rate. Waller and Meehl (1998) provide a technique for classifying cases into putative groups, which enables the calculation of a GFI value.

Simulating Comparison Data

Dimensional comparison data were generated by using factor loadings to reproduce the observed indicator correlation matrix, with the number of factors determined through a preliminary factor analysis of the research data. Each indicator's score distribution was reproduced by using the standard bootstrap technique. Using the research data as an estimate of the population distribution, simulated data contains values sampled randomly, with replacement, from this distribution. Taxonic comparison data were generated by assigning cases to groups and then reproducing indicator correlations and distributions within each group. Ruscio et al. (In press) verified that indicator distributions and correlations are reproduced with good precision and negligible, if any, bias—both within and between groups. Because the analysis of taxonic comparison data generates results representing the pattern to be expected given the fallible criterion used to assign cases to putative groups, it is crucial to choose the criterion variable that best denotes membership status in the putative groups under investigation. In the present study, current diagnoses were used to assign cases to putative MDD+ and MDD– groups. Ten samples each of taxonic and dimensional comparison data were generated for each analysis.

Comparison Curve Fit Index

Calculating the CCFI requires the averaging of each full panel of curves for a given analytic procedure. The root mean square residual of the y values on the averaged curves of the research data and the simulated data is calculated, once to evaluate the fit of the taxonic comparison data, and once to evaluate the fit of the dimensional comparison data:

$$Fit_{RMSR} = \sqrt{\frac{\sum(y_{res.data} - y_{sim.data})^2}{N}},$$

where $y_{res.data}$ refers to a data point on the curve for the research data, $y_{sim.data}$ refers to the corresponding data

point on the curve for simulated taxonic or dimensional data, and N refers to the number of points on each curve. Lower values of Fit_{RMSR} reflect better fit, with perfect fit yielding a value of 0. The two fit values, $Fit_{RMSR\text{-}tax}$ and $Fit_{RMSR\text{-}dim}$, are then integrated into a single index as follows:

$$CCFI = Fit_{RMSR\text{-}dim}/(Fit_{RMSR\text{-}dim} + Fit_{RMSR\text{-}tax})$$

CCFI values can range from 0 to 1, with higher values suggesting better fit for taxonic structure. The index is symmetric about .50 in that this middle value represents equivalent fit for both structures.