Psychological Medicine, 2006, **36**, 973–985. © 2006 Cambridge University Press
doi:10.1017/S0033291706007689 First published online 15 May 2006 Printed in the United Kingdom

A taxometric investigation of unipolar depression in a large community sample

ARI SOLOMON[1]*, JOHN RUSCIO[2], JOHN R. SEELEY[3] AND PETER M. LEWINSOHN[3]

[1] *Department of Psychology, Williams College, Williamstown, MA, USA*; [2] *Department of Psychology, Elizabethtown College, Elizabethtown, PA, USA*; [3] *Oregon Research Institute, Eugene, OR, USA*

ABSTRACT

Background. The question of whether unipolar clinical depression differs categorically from limited depressive complaints has important implications for the disorder's assessment, treatment and research. This crucial issue has proven difficult to resolve, in part because many studies to date have relied on self-report measures or on clinically homogeneous samples. We therefore applied Meehl's taxometric method to a large, clinically heterogeneous sample, and examined the latent structure of depressive episodes using both self-report and structured clinical interview data.

Method. Data were derived from the Oregon Adolescent Depression Project, a large longitudinal community study. All analyses involved more than 1400 participants. MAXEIG (MAXimum EIGenvalue) and base rate estimation were performed separately for Beck Depression Inventory (BDI) items and for DSM-IV-based major depressive episode (MDE) symptoms.

Results. MAXEIG analyses of the BDI and MDE indicator sets appeared to converge on a taxonic structure for unipolar depression. Base rate estimates overall implied a latent depressive episode class that occurs more frequently than diagnosable MDEs but less frequently than persistent depressed or anhedonic mood.

Conclusions. These findings provide tentative support for a categorical conceptualization and make it very clear that the continuity controversy regarding unipolar depression has not yet been decided in favor of dimensionality. To reconcile the conflicting reports to date, several data analytic and sampling issues need to be explored systematically.

INTRODUCTION

Is there any boundary in nature between clinical depression and mere unhappiness? The question has inspired a controversy that has persisted over several decades (e.g. Seligman, 1978; Angst & Merikangas, 1997; Cox *et al.* 1999) with no sign of diminishing in recent years (e.g. Fechner-Bates *et al.* 1994; Flett *et al.* 1997; Cole, 2004). Resolving the controversy could profoundly change how researchers conceptualize the causes, correlates and consequences of depression. For example, unequivocal evidence that episodes of unipolar depression have a distinct etiology, course and/or treatment response, that is that they constitute a clinical taxon (Meehl, 1992), could improve case identification, would challenge the common research practice of allowing subthreshold mood complaints to serve as a proxy for depression, and would challenge theories that assume dimensionality, including Beck's cognitive theory of depression (Clark *et al.* 1999). Conversely, if unipolar depression is dimensional, then replacing current diagnostic criteria with dimensional diagnosis should improve prediction of the course, prognosis and

* Address for correspondence: Dr Ari Solomon, Department of Psychology, Williams College, Williamstown, MA 01267, USA. (Email: asolomon@williams.edu)

outcome of depression, and should enhance statistical power in depression research (Fraley & Waller, 1998).

Striking similarities have been documented between major depressive episodes (MDEs) and so-called *limited*, *subclinical* or *subthreshold* depressive episodes (e.g. Wells *et al.* 1989; Judd *et al.* 1994; Kendler & Gardner, 1998; Lewinsohn *et al.* 2000). But such findings cannot resolve the continuity controversy, because a manifest symptom continuum can conceal a latent clinical taxon (Meehl, 1995). A consensus has therefore emerged that depression researchers must test latent structure using statistical techniques validated for that purpose (Cole, 2004). The taxometric analytic method (Meehl, 1995; Waller & Meehl, 1998), used in the present study, is a leading approach in the study of latent structure (Meehl, 1973; Meehl & Golden, 1982; Meehl & Yonce, 1994, 1996; Waller & Meehl, 1998).

Taxometric studies of clinically selected populations

There have been several taxometric investigations of the structure of depressive symptoms, but most have relied on samples selected on the basis of one or more clinically significant symptoms. For example, Slade and Andrews (2005) reported a dimensional latent structure for depressive symptoms in a very large community sample, but analyses were restricted to participants who had endorsed 2 weeks or longer of pervasive depressed mood and/or anhedonia (i.e. criterion A in the DSM-IV MDE criteria). Prisciandaro and Roberts (2005) inferred dimensionality of depressive symptoms within a large sample of non-patients, all of whom shared a lifetime history of DSM MDE criterion A (persistent depressed or anhedonic mood). Ruscio and Ruscio (2000, Study 1) found evidence for a dimensional latent structure for self-report depressive symptoms in analyses restricted to veterans seeking evaluation for post-traumatic stress disorder, a majority (63%) of whom met criteria for major depressive disorder (MDD). Overall, most studies of selected samples have found evidence for a dimensional structure to depressive complaints, but one important recent exception is Ruscio *et al.* (in press *b*), who found evidence for a clinical depression taxon using a large, diagnostically diverse out-patient psychiatric sample.

A possible weakness of such studies has been implied in several recent reviews of the continuity controversy (e.g. Fechner-Bates *et al.* 1994; Coyne & Schwenk, 1997; Solomon *et al.* 2001; Slade & Andrews, 2005): because individuals with modestly elevated depressive symptoms are included in such studies only if they have pre-existing internalizing psychopathology, their mild depression symptoms may generally represent incipient or partially remitted clinical episodes – *formes frustes* of the same pattern seen in diagnosable cases. By contrast, modestly elevated individuals in the general community could typically exhibit a qualitatively different – benign – symptom pattern, representative of mere transient distress. Indeed, some samples selected for non-affective internalizing psychopathology have nearly universal histories of affective illness (Bleich *et al.* 1997). As depression is conceptualized as a waxing and waning lifetime syndrome (APA, 1994; Judd *et al.* 1998), subthreshold depressive complaints in the context of a history of MDD might usually represent limited recurrences of the clinical symptom pattern. Studies of clinically selected samples might therefore tend to yield dimensional depression solutions because such datasets contain few, if any, true non-taxon cases.

Conversely, several taxometric subtyping studies have reported evidence for subtypes of MDD (e.g. Haslam & Beck, 1994; Ambrosini *et al.* 2002). But such studies typically examine only diagnosable MDD cases, and therefore do not help to resolve the question of continuity/discontinuity across the full range of symptom severity. They cannot resolve, for example, whether very limited endogenous complaints exist along a continuum with pronounced endogenous complaints, or differ categorically (Solomon *et al.* 2001).

Taxometric studies of clinically unselected populations

To our knowledge, four published taxometric reports have explored the latent structure of depressive symptoms using samples that were not uniformly selected for some kind of psychological complaint; two of those studies analyzed self-report [Beck Depression Inventory (BDI) or Minnesota Multiphasic Personality Inventory (MMPI)] depression symptoms. Ruscio & Ruscio (2000, Study 2) reached an inference

of dimensional latent structure for MMPI depression scale scores using a large sample of general medical and psychiatric patients. Ruscio & Ruscio (2002) reached an inference of dimensional latent structure for BDI scores in a large undergraduate sample. Beach and Amir (2003, 2005) inferred a taxonic latent structure for a subset of somatic BDI items in an undergraduate sample, although that interpretation has been controversial (Ruscio *et al.* 2004, Study 1). A fourth study (Hankin *et al.* 2005) found evidence for a dimensional latent structure for depressive symptoms in a large unselected youth sample, but used an interviewer-administered self-report scale [Child and Adolescent Psychopathology Scale (CAPS); Lahey *et al.* 2004] in which each depressive symptom is assessed with reference to its presence at any time in the past year ('during the last twelve months were there times when …'). A high CAPS depression score could therefore represent depression symptoms that occurred during past months or weeks, either separately or simultaneously, and up to 12 months ago. Such patterns would not necessarily capture the core structure of a current MDE, that is the simultaneous current experience of persistent multiple symptoms.

To summarize, on the one hand findings from taxometric studies using clinically unselected samples are somewhat contradictory and limited by exclusive reliance on self-report measures. On the other hand, the implications of dimensional findings in studies of clinically selected samples may be limited because of possible exclusion of most complement class members. Taken together, there appears to be a noteworthy lack of taxometric studies that have examined well-validated structured clinical interview symptom ratings in unselected samples. We believe ours is the first published study to offer such an analysis, and also the first to examine the latent structure of BDI and interview-based MDE symptoms within a single sample.

The present study

Data were drawn from the Oregon Adolescent Depression Project (OADP), a longitudinal epidemiological study of psychopathology in which an initially unselected sample of 1709 high-school students (ages 14–18) was assessed repeatedly to age 30. The OADP sample affords adequate representation of non-depressed, subthreshold, and depressed individuals, and provides both self-report and reliable structured symptom ratings. Previous OADP publications have reported on the prevalence and incidence of mental disorders, the correlates of diagnosable and subthreshold depression, and the stability of psychopathology into adulthood. This is the first OADP report on the latent structure of unipolar depression.

METHOD

Participants

The original OADP sample consisted of 1709 students drawn at random from a total population of 10 200 students at nine high schools in western Oregon; these students were paid to complete a psychological assessment and diagnostic interview (T1). One year later (T2), 1507 of the participants were assessed again, with repeat and additional instruments. Prior publications have described the local representativeness of the T1/T2 OADP sample and have compared participating and non-participating students (e.g. Lewinsohn *et al.* 1993). Later waves were collected, but our analyses focused on the T1/T2 waves for which there were essentially complete repeated administrations of depression measures.

Depression measures

Diagnostic interviews

The T1 and T2 Axis I diagnostic interview was based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS; Orvaschel *et al.* 1982), modified to assess lifetime and current disorders according to DSM-III-R criteria (APA, 1987). MDD diagnoses at T1 and T2 conformed with both DSM-III-R and DSM-IV criteria (APA, 1994). K-SADS interviews at T2 were combined with the Longitudinal Interval Follow-up Evaluation (LIFE; Keller *et al.* 1987), and the combined instrument was modified to probe for continuing or emergent episodes since the previous interview. Diagnostic interviewers were carefully selected, trained and tested, and reliability of diagnoses was very good to excellent (see Lewinsohn *et al.* 1993 for detailed descriptions of rater qualifications and reliabilities). For current MDE, κ was 0·88 at T1 and 0·83 at T2.

Hamilton Depression Rating Scale items

The K-SADS/LIFE interview module was augmented with a 14-item version of the Hamilton Depression Rating Scale (HAMD; Hamilton, 1960). The original HAMD is a 21-item clinician-rated measure of depressive symptoms that is completed subsequent to a detailed clinical interview. Each HAMD symptom item is scaled (0–3 or 0–4) so that higher scores represent more severe symptoms. Some HAMD items closely overlap with DSM MDE symptoms, but the scaling is substantially different. For example, the DSM requires only a judgment of clinically significant/non-significant for psychomotor retardation, and permits consideration of unobserved retardation. By contrast, the HAMD focuses exclusively on retardation observed during the interview and requires a more fine-grained differentiation ('slight retardation at interview', 'obvious retardation', 'interview difficult', 'complete stupor'). For our analyses only the 14 HAMD items with the highest validity coefficients were considered for inclusion in composite DSM/HAMD indicators, and of those, only the ones that assessed MDE symptoms were retained. Rater agreement statistics for this HAMD short form are unavailable but are presumed to be adequate given the excellent reliabilities of the K-SADS and LIFE interview modules within which the items were rated.

Beck Depression Inventory (BDI)

All participants completed the BDI as part of a larger questionnaire battery at T1 and T2. The BDI exhibits high validity as an index of symptom change in diagnosed depressives (Beck *et al.* 1988), but may be valid only as a measure of distress, dysphoria or negative affectivity in non-clinical cases (Kendall *et al.* 1987). Each of the 21 BDI items estimates the severity of a depressive symptom during the past week on a scale of 0 (completely absent) to 3 (most severe). In the present sample BDI scores were positively skewed at both T1 (mean = 7·05, S.D. = 7·48, skewness = 1·66, $\alpha = 0{\cdot}88$, $n = 1709$) and T2 (mean = 5·41, S.D. = 7·13, skewness = 2·27, $\alpha = 0{\cdot}90$; $n = 1482$). BDI descriptives in the present sample resemble the values reported for undergraduate samples (e.g. mean = 7·42, S.D. = 6·67, $\alpha = 0{\cdot}89$ in Dozois *et al.* 1998; mean = 6·37, S.D. = 6·26, $\alpha = 0{\cdot}87$ in Ruscio & Ruscio, 2002) and for high-school samples (e.g. mean = 8·45, S.D. = 7·41 in Ehrenberg *et al.* 1990).

Exclusionary criteria

Participants were omitted from analyses if they met any of five DSM-IV MDD exclusionary criteria, each of which represents a possible clinical taxon that co-occurs with depression and could therefore confound interpretation of the latent structure of depression. These exclusions are: lifetime bipolar diagnosis; lifetime non-affective psychotic diagnosis; current bereavement; current organic mood syndrome; or current substance-induced mood syndrome. This resulted in 44 exclusions at T1 (2·4%) and 39 at T2 (2·3%).

Taxometric analyses

The taxometric method is designed to determine whether the structural relationships among indicators of a particular construct (e.g. depression) imply the existence or non-existence of two latent groups: a taxon (e.g. depressed individuals) and its complement (e.g. non-depressed individuals). Instead of testing a directional or null hypothesis, consistent evidence of either taxonic or dimensional structure is sought across multiple tests. Confidence increases when multiple analyses converge on a taxonic or dimensional structure.

Indicator sets

Multiple measures of depression and waves of data were available. The primary selection criteria were that each indicator should be able to distinguish a putative taxon and complement class with sufficient validity, and that pairs of indicators be correlated substantially more highly in the full sample than within putative groups. Our approach was to analyze many candidate indicator sets and retain those that met these requirements and yielded taxometric results that were interpretable (i.e. more consistent with either taxonic or dimensional structure) rather than ambiguous (i.e. equally consistent with both of these structural models).

BDI item pairs in configurations developed and described by Ruscio & Ruscio (2000) were used (1654 cases at T1 and 1424 at T2).†

† Following Ruscio & Ruscio (2000), we also constructed indicator sets using individual (rather than paired) BDI items from both T1 and T2 data. Analyses of those additional indicators were consistent with the paired-item analyses and are omitted here to conserve space.

Table 1. *Items used to form DSM symptom–Hamilton composite indicator sets*

Composite	DSM symptoms	Hamilton items
1 Insomnia	4 (sleep disturbance)	4 (Insomnia)
2 Appetite/weight	3a (increased appetite) 3b (decreased appetite)	16 (loss of weight)
3 Mood disturbance	1 (depressed mood) 2 (loss of interest)	1 (depressed mood)
4 Suicidality	9 (suicidality)	3 (suicide)
5 Fatigue/anergia	5a (psychomotor retardation) 6 (fatigue)	8 (retardation)
6 Guilt/worthlessness	7 (guilt/worthlessness) 24 (worthlessness)	2 (feelings of guilt)
7 Cognitive impairment	8a (impaired thinking/concentration) 8b (indecision)	
8 Agitation	5b (psychomotor agitation)	9 (agitation)

Table 2. *Indicator validity and within-group correlations for BDI and DSM symptom–Hamilton indicator sets*

			Within-group correlations	
	Indicator validity (Cohen's *d*)		Taxon	Complement
Indicator set	Range	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)
Permissive criterion				
BDI, T1	0·68–1·25	1·00 (0·24)	0·42 (0·09)	0·46 (0·05)
BDI, T2	0·91–1·46	1·22 (0·27)	0·49 (0·10)	0·47 (0·05)
DSM-Hamilton	0·82–6·28	1·95 (1·77)	0·28 (0·07)	0·16 (0·09)
Restrictive criterion				
BDI, T1	0·98–1·79	1·37 (0·34)	0·31 (0·10)	0·49 (0·05)
BDI, T2	1·01–1·80	1·42 (0·35)	0·54 (0·09)	0·49 (0·06)
DSM-Hamilton	2·78–4·15	3·27 (0·58)	0·08 (0·16)	0·23 (0·12)

Using a permissive criterion (currently meets DSM criterion A of persistent depressed mood or anhedonia in structured diagnostic interview), the base rate for the putative current depression taxon was 0·120 at T1 and 0·08 at T2; using a restrictive criterion (currently diagnosed MDD), the base rate was 0·024 at T1 and 0·032 at T2. Even the lower estimates of the putative taxon's base rate were judged adequate for taxometric analyses in such a large sample.

A combination of T1 Hamilton items and T1 DSM symptoms yielded two additional indicator sets, each of which contained eight composite indicators (see Table 1). Complete composite indicator data were available for 1629 individuals; using the permissive criterion the base rate was 0·120, and using the restrictive criterion it was 0·024. Estimates of indicator validity and within-group correlations for all indicator sets retained for analysis are shown in Table 2.

Analysis plan

We performed three taxometric procedures using each set of indicators: MAMBAC (Mean Above Minus Below A Cut; Meehl & Yonce, 1996), MAXEIG (MAXimum EIGenvalue; Waller & Meehl, 1998), and L-Mode (Latent Mode; Waller & Meehl, 1998). Perhaps because the putative taxon was so small, neither MAMBAC nor L-Mode yielded interpretable results. All of the MAMBAC curves sloped upward at the upper end, a pattern equally consistent with a small taxon or a latent dimension with positively skewed indicators (Ruscio *et al.* 2004). All of the L-Mode graphs were ambiguous in the sense that perturbations toward the right end of each frequency distribution could either be viewed as chance-level occurrences,

consistent with the unimodal shape expected for dimensional structure, or interpreted as bumps representing the emergence of a second – and much smaller – group, consistent with the bimodal shape expected for taxonic structure.

Provided that all indicators are sufficiently valid, the multivariate MAXEIG procedure can be a more powerful tool than MAMBAC and L-Mode for determining whether a small putative taxon is present; the MAXEIG inchworm consistency test (Waller & Meehl, 1998) helps to resolve otherwise ambiguous results. This technique involves systematically increasing the number of overlapping windows in the analysis. A small taxon should produce increasingly peaked curves with larger numbers of windows, whereas curves for dimensional data should remain cusped (non-peaked) even with the largest number of windows. In the present investigation, the inchworm consistency test proved invaluable because of the small size of the putative taxon.

Because the DSM-Hamilton indicator sets were constructed based on a conceptual aggregation of items, nuisance covariance may have prevented detection of a taxon if one existed. Two techniques were therefore used to pare the full set of indicators down to smaller numbers that represented more conceptually and empirically distinct facets of depressive episodes, with the goal of reducing nuisance covariance to tolerable levels. First, we began MAXEIG analyses with all eight indicators but successively dropped a single indicator with the lowest indicator validity and highest nuisance covariance until only three indicators remained. These analyses are referred to as the 'best indicators' analyses. Second, we combined one indicator at a time for successive analyses until only three composite indicators remained. (The item with the lowest difference in its average full-sample and within-groups correlations with all other indicators was selected to be collapsed with another item. The latter item was identified as the one whose correlational difference score, that is full-sample minus within-groups, would be most improved by combining it with the former item.) We refer to those analyses as the 'combined indicators' series. Finally, several artificial dimensional data sets were created using a latent factor model and eight indicators ($n = 1500$), and was submitted to the best-indicator and combined-indicator algorithms. None of the MAXEIG inchworm consistency tests yielded a peaked curve. Therefore, both techniques appear to be resistant to spuriously taxonic results.

RESULTS

BDI indicator sets

MAXEIG results for T1 and T2 paired-item BDI indicator sets are shown in Fig. 1. Full panels of curves were inspected to make sure that the averaged curves in Fig. 1 fairly represented a trend observed across all indicators. The paired-item indicators exhibit the prototypical shape for a small taxon clearly for the T2 data, and somewhat less clearly for the T1 data, although both panels exhibit a progression from initial ambiguous right-end cusps toward more fully defined taxonic peaks.

Consistent with an inference of taxonic structure, taxon base rate estimates remained highly consistent as the number of windows increased within each application of the inchworm consistency test. As Ruscio *et al.* (2004) noted, such consistency in the presence of a small taxon would be expected, whereas when indicators represent a dimensional construct base rate estimates would be expected to fall with increasing numbers of windows. At T1, paired-item base rate estimates ranged only from 0·083 to 0·087. At T2, paired-item estimates ranged from only 0·099 to 0·104.

DSM-Hamilton indicator sets

Results for these analyses are shown in Fig. 2(*a*, *b*). A fairly consistent taxonic pattern of curves emerged in the combined indicators MAXEIG series; curves initially rose ambiguously toward a right-end cusp, but as the number of indicators was reduced the inchworm consistency test revealed better-defined taxonic peaks. By the time the indicator set had been reduced to five indicators, results were consistent with expectations for a small taxon. Reducing the number of indicators beyond this point yielded ambiguous results, possibly because there were too few indicators of sufficient validity to distinguish taxonic from dimensional structure. Consistent with that possibility, in the best-indicators analyses – which retained the three most valid indicators – a large majority of the curves continued to show a taxonic peak even when the number of indicators had been reduced to 3.

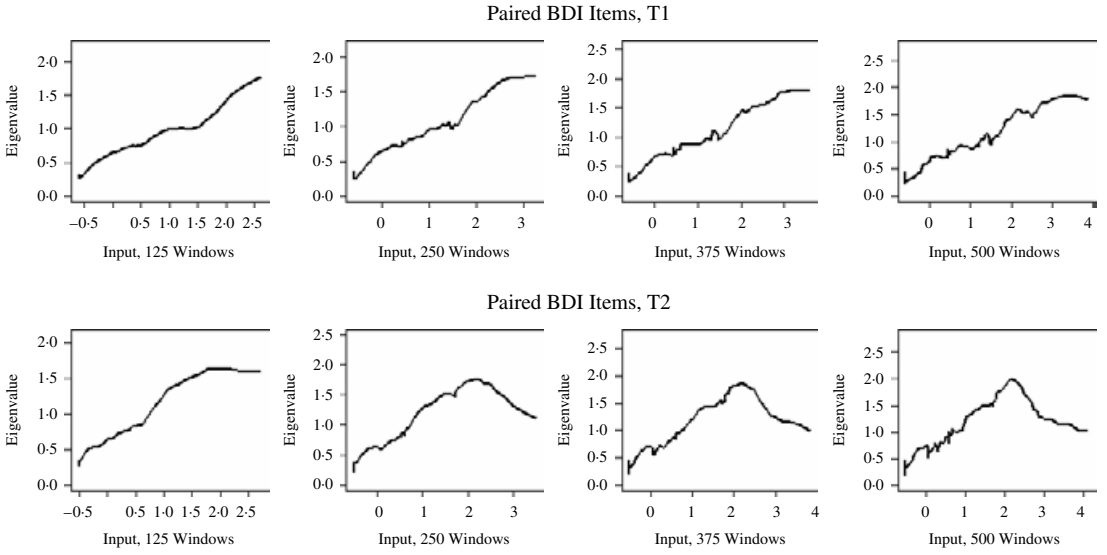


FIG. 1. MAXEIG results for analyses of BDI indicator sets.

DSM-Hamilton taxon base rate estimates were consistent with one another within and across all estimates (for the 24 best-indicator analyses: mean = 0·065, S.D. = 0·011; for the 24 combined-indicator analyses: mean = 0·053, S.D. = 0·006). All base rate estimates fell within the range of plausible values bounded at the lower end by the restrictive and permissive depression class criteria described above. The base rate estimates are lower in these analyses than in the BDI analyses, but this is consistent with the higher within-group indicator correlations in the BDI data (a relationship documented in Ruscio *et al.* in press *a*). The full sequence of taxometric analyses was re-run after excluding all cases matching the DSM-IV melancholic MDE specifier. The taxon base rate estimates were somewhat reduced, as expected, but the pattern of results was unchanged.

DISCUSSION

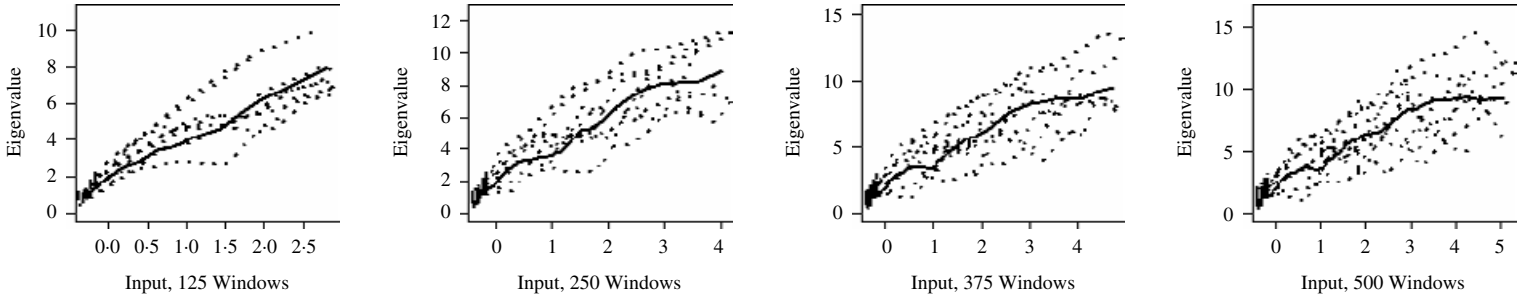
Much is at stake in the controversy over the latent structure of depressive episodes; the outcome of this issue could alter how mood symptoms are diagnosed, studied and treated. For example, if depressive episodes prove to be taxonic, then intensive study of taxon members could lead to a better ability to differentiate clinically meaningful presentations from transient distress. If depressive episodes prove to be dimensional, then categorical decision models may need to be replaced with assessment strategies that emphasize the importance of symptom severity as a continuous variable.

In the present study we used the taxometric method to examine the episode continuity question, relying on expert clinical ratings and a large, clinically unselected sample. Evidence for taxonic structure was found using the inchworm consistency test, an approach to MAXEIG analyses that can help to distinguish a small latent taxon from a dimensional construct with positively skewed indicators (Ruscio *et al.* 2004). The taxonic pattern emerged fairly consistently in MAXEIG analyses across both self-report and clinical symptom indicator sets, and the pattern persisted even when cases of melancholic depression were excluded.

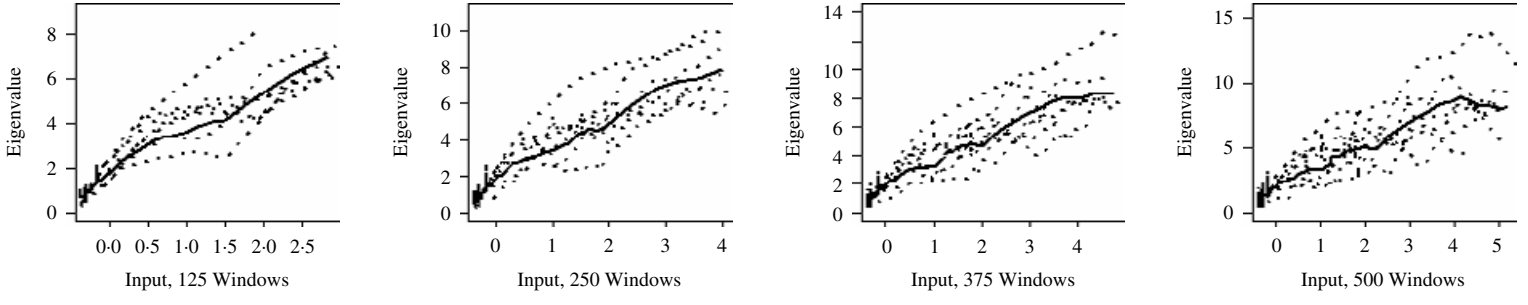
To our knowledge this is the first published taxometric investigation of depression to rely on both a predominantly unselected sample and DSM-defined depressive symptoms; and the first to compare the latent structure of self-report and expertly rated symptoms within a single sample. Our primary finding contradicts dimensional findings from several studies of selected samples. However, as noted earlier, it is not yet clear whether studies of selected samples can resolve whether a structural boundary exists

(*a*) DSM-Hamilton Combined Indicators, T1

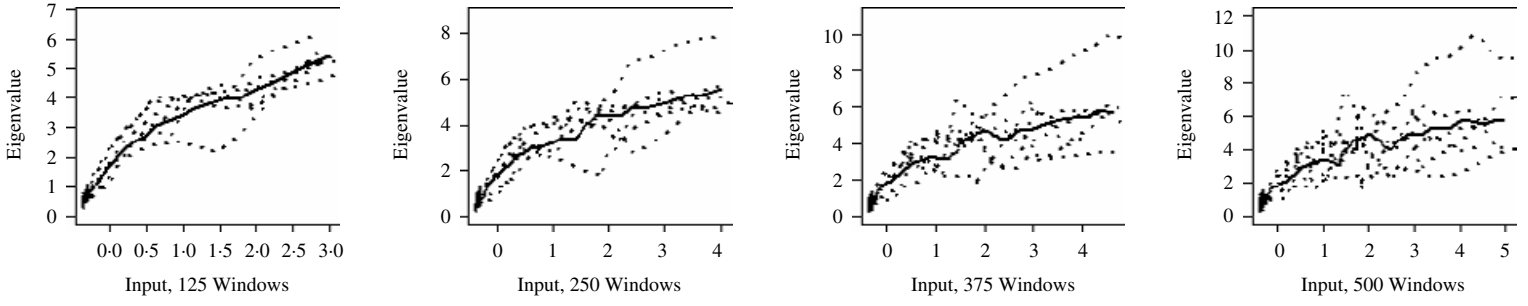
8 Indicators



7 Indicators (#8 combined with #1)



6 Indicators (#4 combined with #1/8)



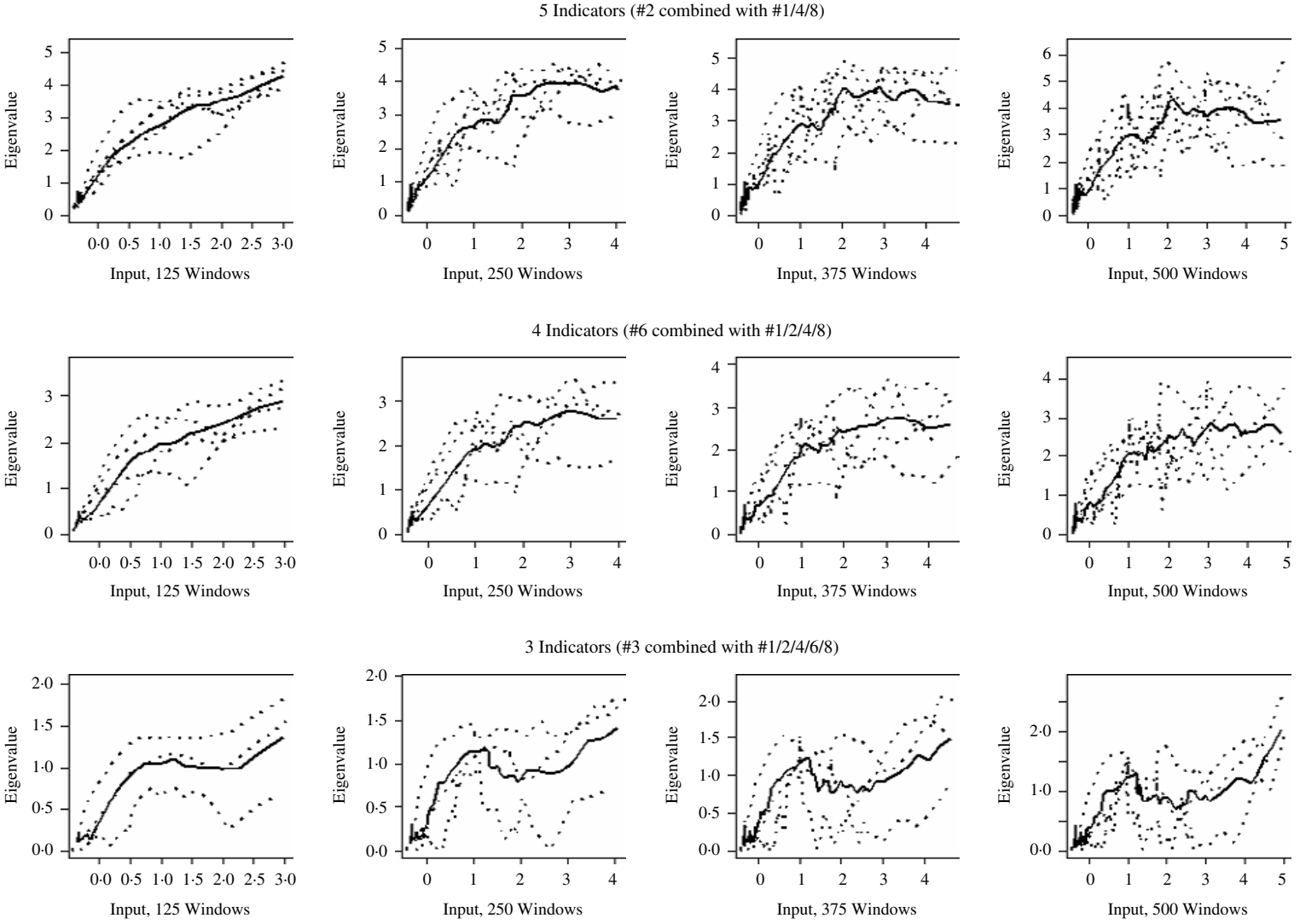
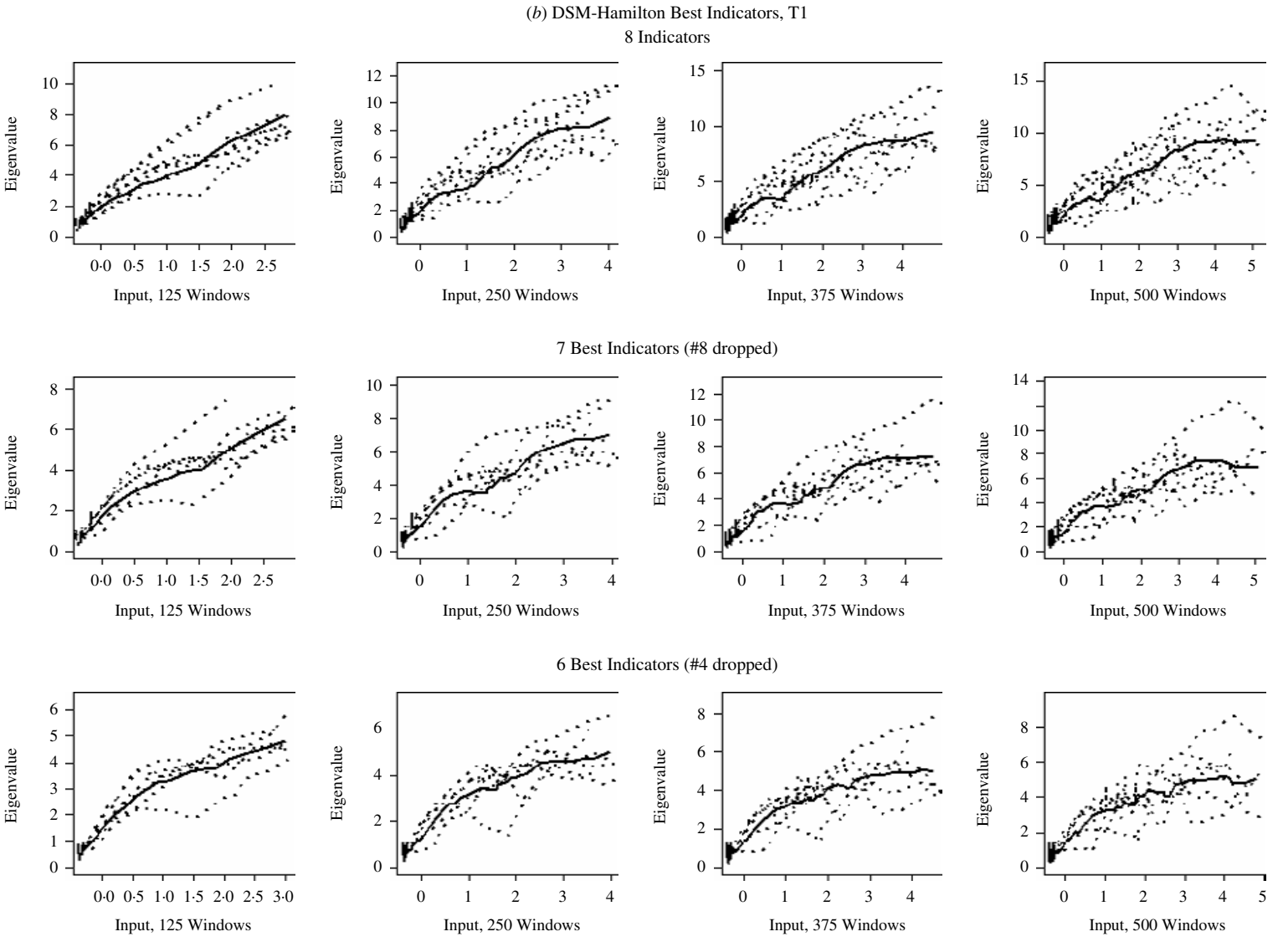


FIG. 2(*a*). MAXEIG results for analyses of DSM-Hamilton 'combined-item' indicator sets using the T1 data.


(b) DSM-Hamilton Best Indicators, T1
8 Indicators
7 Best Indicators (#8 dropped)
6 Best Indicators (#4 dropped)
Eigenvalue
Input, 125 Windows
Input, 250 Windows
Input, 375 Windows
Input, 500 Windows

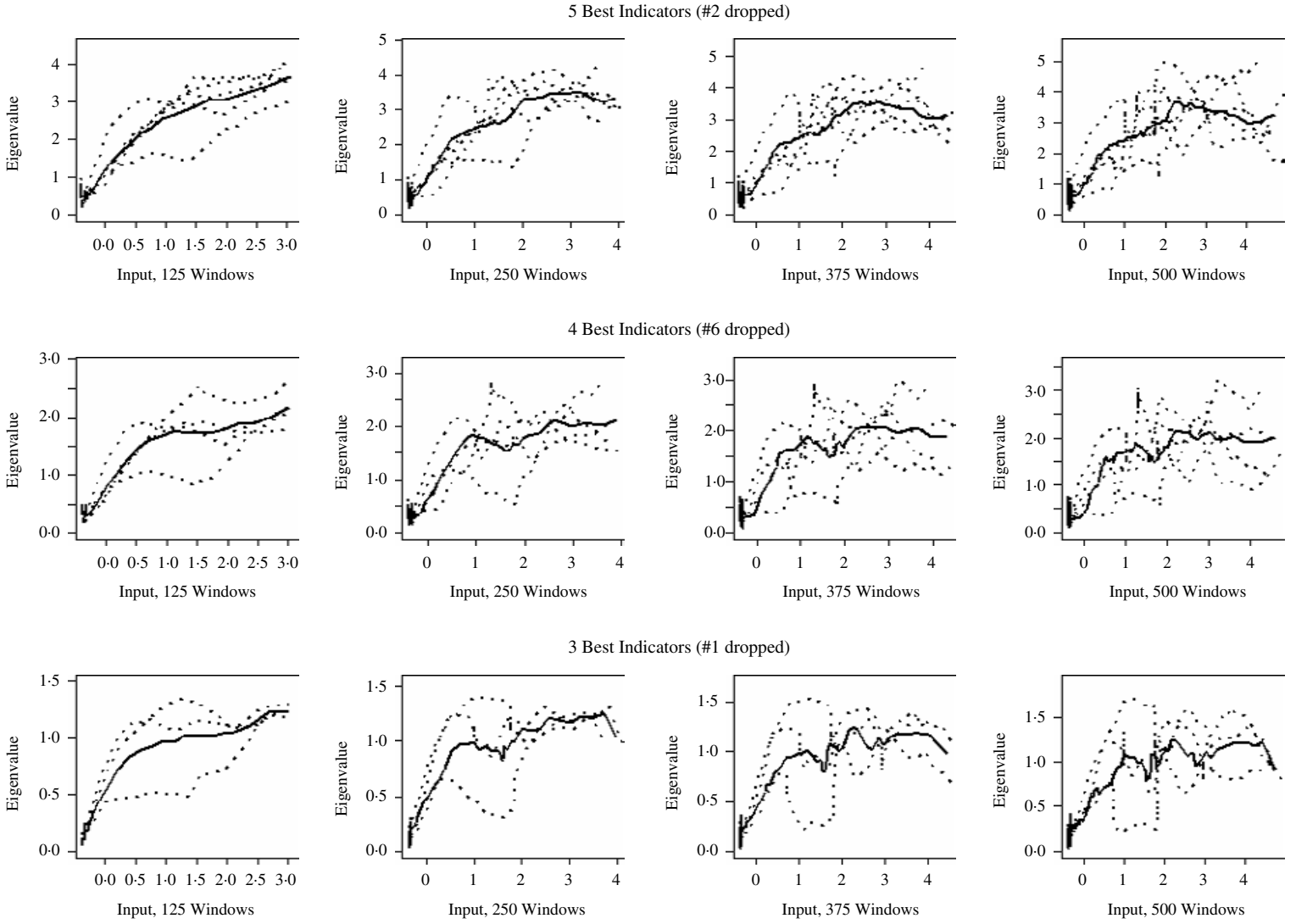


FIG. 2(*b*). MAXEIG results for analyses of DSM-Hamilton 'best item' indicator sets using the T1 data.

in the general population. Furthermore, our conclusions are quite consistent with a recent finding of taxonic latent structure in a large, diagnostically heterogeneous out-patient sample (Ruscio *et al*. in press *b*).

Despite the study's distinctive strengths, four limitations merit particular consideration. First, although the preponderance of MAXEIG curves appeared consistent with a taxonic structure, one of two subsets of curves (the 'combined indicators' set) became ambiguous when fewer than five composite indicators were used in analyses. Second, only one taxometric procedure yielded interpretable results, probably due to the challenge of detecting a small taxon. Although the MAXEIG results were interpretable as taxonic across multiple indicator sets, in different configurations, and with different numbers of overlapping windows, consistency of results across multiple procedures would have afforded greater confidence in the conclusion of taxonic structure. Third, our findings so far bear only on the structure of adolescent depression. Fourth, establishing evidence that a taxon exists is only a first step; determining the clinical implications of taxon membership requires additional investigation.

For all of those reasons it is important to regard these inferences of taxonic structure as fairly tentative, and to seek their replication and extension in other large, unselected samples. Until then, any firm conclusions about the structure of depressive episodes or the distinguishing characteristics of taxon members strike us as premature, particularly in light of conflicting findings in methodologically diverse studies. What we can say unequivocally is that both our present findings and another recent report of taxonic structure in an interviewed sample (Ruscio *et al*. in press *b*) contrast sharply with suggestions that the latent structure of episodic depression has been resolved by recent studies in favor of dimensionality (e.g. Schmidt *et al*. 2004; Slade & Andrews, 2005). We would argue instead that this perdurable controversy – now well into its third decade – still resists a final verdict.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute of Mental Health (RO3 MH067961), awarded to A.S.

DECLARATION OF INTEREST

None.

REFERENCES

Ambrosini, P., Bennett, D. S., Cleland, C. M. & Haslam, N. (2002). Taxonicity of adolescent melancholia: a categorical or dimensional construct? *Journal of Psychiatric Research* **36**, 247–256.

Angst, J. & Merikangas, K. (1997). The depressive spectrum: diagnostic classification and course. *Journal of Affective Disorders* **45**, 31–40.

APA (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd revised edn). American Psychiatric Press: Washington, DC.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn). American Psychiatric Press: Washington, DC.

Beach, S. R. H. & Amir, N. (2003). Is depression taxonic, dimensional, or both? *Journal of Abnormal Psychology* **112**, 228–236.

Beach, S. R. H. & Amir, N. (2005). Self-reported depression is taxonic. *Journal of Psychopathology and Behavioral Assessment*. doi:10.1007s10862-005-9007-4. Published online 2 December 2005.

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

Bleich, A., Koslowsky, M., Dolev, A. & Lerer, B. (1997). Post-traumatic stress disorder and depression: an analysis of comorbidity. *British Journal of Psychiatry* **170**, 479–482.

Clark, D. A., Beck, A. T. & Alford, B. A. (1999). *Scientific Foundations of Cognitive Theory and Therapy of Depression*. Wiley: New York.

Cole, D. A. (2004). Taxometrics in psychopathology research: an introduction to some of the procedures and related methodological issues. *Journal of Abnormal Psychology* **113**, 3–9.

Cox, B. J., Enns, M. W., Borger, S. C. & Parker, J. D. A. (1999). The nature of the depressive experience in analogue and clinically depressed samples. *Behaviour Research and Therapy* **37**, 15–24.

Coyne, J. C. & Schwenk, T. L. (1997). The relationship of distress to mood disturbance in primary care and psychiatric populations. *Journal of Consulting and Clinical Psychology* **65**, 161–168.

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

Ehrenberg, M. F., Cox, D. N. & Koopman, R. F. (1990). The prevalence of depression in high school students. *Adolescence* **25**, 905–912.

Fechner-Bates, S., Coyne, J. C. & Schwenk, T. L. (1994). The relationship of self-reported distress to depressive disorders and other psychopathology. *Journal of Consulting and Clinical Psychology* **62**, 550–559.

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

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

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* **23**, 56–61.

Hankin, B. L., Fraley, R. C., Lahey, B. B. & Waldman, I. D. (2005). Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *Journal of Abnormal Psychology* **114**, 96–110.

Haslam, N. & Beck, A. T. (1994). Subtyping major depression: a taxometric analysis. *Journal of Abnormal Psychology* **103**, 686–692.

Judd, L. L., Akiskal, H. S., Maser, J. D., Zeller, P. J., Endicott, J., Coryell, W., Paulus, M. P., Kunovac, J. L., Leon, A. C.,

Mueller, T. I., Rice, J. A. & Keller, M. B. (1998). A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Archives of General Psychiatry* **55**, 694–700.

Judd, L. L., Rapaport, M. H. & Paulus, M. P. (1994). Subsyndromal symptomatic depression (SSD): a new mood disorder? *Journal of Clinical Psychiatry* **55**, 18–28.

Keller, M. B., Lavori, P. W., Friedman, B. & Nielsen, E. (1987). The longitudinal interval follow-up evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* **44**, 540–548.

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

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

Lahey, B. B., Applegate, B., Waldman, I. D., Loft, J. D., Hankin, B. L. & Rick, J. (2004). The structure of child and adolescent psychopathology: generating new hypotheses. *Journal of Abnormal Psychology* **113**, 358–385.

Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R. & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology* **102**, 133–144.

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

Meehl, P. E. (1973). MAXCOV-HITMAX: a taxometric search method for loose genetic syndromes. In *Psychodiagnosis: Selected Papers* (ed. P. E. Meehl), pp. 200–224. University of Minnesota Press: Minneapolis.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality* **60**, 117–174.

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

Meehl, P. E. & Golden, R. R. (1982). Taxometric methods. In *Handbook of Research Methods in Clinical Psychology* (ed. P. C. Kendall and J. N. Butcher), pp. 127–181. Wiley: New York.

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

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

Orvaschel, H., Puig-Antich, J., Chambers, W., Tabrizi, M. A. & Johnson, R. (1982). Retrospective assessment of child psychopathology with the Kiddie SADS-E. *Journal of the American Academy of Child Psychiatry* **21**, 392–397.

Prisciandaro, J. J. & Roberts, J. E. (2005). A taxometric investigation of unipolar depression in the National Comorbidity Survey. *Journal of Abnormal Psychology* **114**, 718–728.

Ruscio, A. M. & Ruscio, J. (2002). The latent structure of analogue depression: should the Beck Depression Inventory be used to classify groups? *Psychological Assessment* **14**, 135–145.

Ruscio, J., Haslam, N. & Ruscio, A. M. (in press *a*). *Introduction to the Taxometric Method: A Practical Guide*. Lawrence Erlbaum Associates: Mahwah, NJ.

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

Ruscio, J., Ruscio, A. M. & Keane, T. M. (2004). Using taxometric analysis to distinguish a small latent taxon from a latent dimension with positively skewed indicators: the case of involuntary defeat syndrome. *Journal of Abnormal Psychology* **113**, 145–154.

Ruscio, J., Zimmerman, M., McGlinchey, J. B., Chelminsky, I. & Young, D. (in press *b*). Diagnosing major depressive disorder: XI. A taxometric investigation of the categorical-dimensional debate on the structure underlying DSM-IV symptoms. *Journal of Nervous and Mental Disease*.

Schmidt, N. B., Kotov, R. & Joiner, T. E. (2004). *Taxometrics: Toward a New Diagnostic Scheme for Psychopathology*. American Psychological Association: Washington, DC.

Seligman, M. E. (1978). Comment and integration. *Journal of Abnormal Psychology* **87**, 165–179.

Slade, T. & Andrews, G. (2005). Latent structure of depression in a community sample: a taxometric analysis. *Psychological Medicine* **35**, 489–497.

Solomon, A., Haaga, D. A. & Arnow, B. A. (2001). Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research. *Journal of Nervous and Mental Disease* **189**, 498–506.

Waller, N. G. & Meehl, P. E. (1998). *Multivariate Taxometric Procedures: Distinguishing Types from Continua*. Sage: Thousand Oaks, CA.

Wells, K. B., Stewart, A., Hays, R. D., Burnham, M. A., Rogers, W., Daniels, M., Berry, S., Greenfield, W. & Ware, J. (1989). The functioning and wellbeing of depressed patients: results from the Medical Outcomes Study. *Journal of the American Medical Association* **262**, 914–919.