Published in final edited form as:

Psychol Aging. 2010 March ; 25(1): 142–156. doi:10.1037/a0018514.

The structure of late-life depressive symptoms across a 20 year span: A taxometric investigation

Jason M. Holland, Kathleen K. Schutte, Penny L. Brennan, and Rudolf H. Moos

Center for Health Care Evaluation, VA Palo Alto Health Care System, Stanford University Medical Center

Abstract

Past studies of the underlying structure of depressive symptoms have yielded mixed results, with some studies supporting a continuous conceptualization and others supporting a categorical one. However, no study has examined this research question with an exclusively older adult sample, despite the potential uniqueness of late-life depressive symptoms. In the present study, the underlying structure of late-life depressive symptoms was examined among a sample of 1289 individuals across three waves of data collection spanning 20 years. A taxometric methodology was employed using indicators of depression derived from the Research Diagnostic Criteria. Maximum eigenvalue (MAXEIG) analyses and inchworm consistency tests generally supported a categorical conceptualization and identified a group that was primarily characterized by thoughts about death/suicide. However, compared to a categorical depression variable, depressive symptoms treated continuously were generally better predictors of relevant criterion variables. These findings suggest that thoughts of death and suicide may characterize a specific type of late-life depression, yet a continuous conceptualization still typically maximizes the predictive utility of late-life depressive symptoms.

Keywords

Depression; Taxometric Method; Older Adults; Suicidal Ideation; Aging

The issue of whether depressive symptoms should be conceptualized as discrete categories (e.g., depressed vs. non-depressed individuals) or as a smooth continuum has been debated for decades (Seligman, 1978; Solomon, Ruscio, Seeley, & Lewinsohn, 2006). Theoretical as well as practical concerns, such as those related to insurance billing, have influenced this debate. Empirical efforts have also been attempted with mixed-age or exclusively younger adult samples that aim to objectively test hypotheses about the underlying structure of depressive symptoms (i.e., as either continuous or categorical; Haslam, 2003). In general, these studies have yielded mixed results, with most studies supporting a continuous structure (see Haslam, 2003, for a review) and a handful of other studies supporting a categorical structure (Beach & Amir, 2006; Ruscio, Brown, & Ruscio, in press; Ruscio, Zimmerman, McGlinchey, Chelminski, & Young, 2007; Solomon et al., 2006). Because this issue has yet to be investigated

Direct correspondence to: Jason M. Holland, Center for Health Care Evaluation (152-MPD), VA Palo Alto Health Care System, 795 Willow Rd., Menlo Park, CA 94025. Phone: (650) 493-5000 ext. 23426. jmholland@gmail.com.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/pag

with an older adult sample, the present study aims to enhance understanding of late-life depressive symptoms by investigating the underlying structure of these symptoms. To examine the consistency of this structure, we explore this research question across three waves of data spanning 20 years.

Research evidence suggests that depressive symptoms among older adults differ from those of younger individuals in important ways. In particular, some studies have indicated that depressed older adults are less likely than depressed younger adults to endorse affective symptoms of depression (e.g., feeling sad) and often present with somatic complaints (e.g., fatigue, psychomotor retardation) with no obvious medical etiology (Gallo, Anthony, & Muthen, 1994; Gallo, Rabins, & Anthony, 1999; Gottfries, 1998). Past research has also shown that certain causal factors, such as vascular changes and acute late-life stressors (e.g., bereavement, health problems, retirement), may play a more prominent role during late-life (Alexopoulos et al., 1997; Parker, Hadzi-Pavlovic, Mitchell, & Wilhem, 2003; Van den Berg et al., 2001).

Given these dissimilarities, it seems possible that the underlying structure of depressive symptoms might differ for older adults. For example, it has been suggested that affective symptoms of depression, which are reported less frequently by older adults (Gallo et al., 1994), may appear more continuous compared to depression-related somatic complaints (e.g., related to appetite, sleep, sex drive), which may operate in more of an all-or-nothing type of fashion (Beach & Amir, 2003; Gilbert, 2000). Proponents of this view reason that somatic complaints represent the disruption of various circadian rhythms, and that these disruptions might occur together, producing a dramatic shift in one's self presentation. Indeed, there is some preliminary evidence supporting the notion that somatic symptoms of depression are categorical, and emotional and cognitive components of depression are more continuous (Beach & Amir, 2003). Such findings might suggest that a continuous distribution of late-life depressive symptoms would be unlikely. However, a conclusion of this kind would conflict with the majority of mixed-age studies, which have generally favored a continuous conceptualization (Haslam, 2003).

Considering the ambiguity about whether late-life depressive symptoms are distributed continuously or categorically, the primary purpose of the present study is to examine empirically the latent structure of these symptoms among a large group of community dwelling older adults across three time points, from ages 55–65 (baseline) to 65–75 (10-year follow-up) and 75–85 (20-year follow-up). To address this issue, we will use taxometric methodology, a set of statistical procedures that are specifically designed to distinguish between a categorical structure (which permits dimensional variation within categories) and a continuous structure (Meehl, 1995). This methodology is well suited for our research question and overcomes many of the limitations of other statistical approaches that have been used for this purpose, such as cluster analysis, finite mixture modeling, or latent class analysis. These other approaches have been criticized primarily for their empirical and conceptual limitations in identifying the number of categories that make up a given construct—limitations that stem largely from reliance on a single set of indicators to identify multiple categorical boundaries (Ruscio, Haslam, & Ruscio, 2006).

In contrast to these other strategies, taxometric methodology is typically used to search for a single categorical boundary. Additionally, rather than simply searching for clusters or categories in the data, this method rigorously tests two competing hypotheses (i.e., categorical vs. continuous) about the latent structure of a construct at the same time (Ruscio et al., 2006). Given these unique strengths, taxometric methodology is used in this study to test two competing hypotheses, namely that (1) depressive symptoms among older adults are distributed categorically (i.e., depressed vs. non-depressed individuals) or (2) late-life depressive

symptoms are distributed continuously. These competing hypotheses are examined at all three waves of data collection, in order to provide a consistency test (a hallmark of the taxometric method) that might strengthen the generalizability of results across a wide spectrum of age ranges in later-life.

In this study, we also consider the relative utility of conceptualizing late-life depressive symptoms as continuous or dichotomous with regard to predicting several relevant outcomes. If a construct is dichotomous and the variability within classes mainly represents measurement error, statistical power should not be significantly diminished when this “noise” is removed and the construct is treated categorically—an exception to the conventional wisdom that categorizing a construct measured on a dimensional scale always results in reduced power. Alternately, if symptoms are distributed continuously or if variability within latent categories is meaningful, a continuous measure of depressive symptoms might be more closely associated with relevant criteria. Late-life depressive symptoms have been shown to be associated with medical problems, problem drinking, antidepressant use, and professional help seeking (Atkinson, 1999; Grunebaum, Oquendo, & Manly, 2008; Parslow & Jorm, 2000; Zeiss, Lewinsohn, & Rohde, 1996). Accordingly, each of these factors served as criterion variables in the present study.

Method

Participants and Recruitment

After obtaining informed consent, participants were initially recruited as part of a larger sample of 1884 individuals who were between the ages of 55 and 65 and had contact with a health care facility within the previous 3 years. Because the overall project for which these participants were recruited was designed to focus on late-life drinking behavior, lifetime non-drinkers were, at the recruitment phase, excluded from this sample. However, it is similar to other community based samples of late-middle aged and older adults with regard to health characteristics, such as the presence of chronic illness and rate of hospitalization (Brennan & Moos, 1990). Participants provided demographic information and information about depressive symptoms, medical problems, medication usage, and professional help seeking at baseline and follow-up assessments 10- and 20-years later (see Brennan & Moos, 1990; Moos, Schutte, Brennan, & Moos, 2009, for additional details).

The subset of 1289 participants involved in the present study provided data at both the baseline and 10-year follow-up assessments. Thus, in this investigation the baseline and 10-year follow-up were made up of the same individuals, which allowed us to observe the degree of consistency in our results across these two waves. Of the original 1884 participants surveyed at baseline, 489 had died by the 10-year follow-up, and 1291 individuals (92%) of the 1395 individuals who remained alive participated at this time point. Two individuals were removed due to missing information about their depressive symptoms. The sample was predominantly Caucasian (92%), and a majority of participants were men (59%). At baseline, the average age was 61.3 years ($SD = 3.2$). About half (51%) of participants were employed and 71% were married. The average income was approximately \$40,000/year, and participants had, on average, completed a little over 14 years of education. At baseline, most (89%) participants were currently consuming alcohol, and 63% of participants had experienced no alcohol-related problems during the past year. At the time of the 20-year follow-up, an additional 480 participants had died, and 76 individuals were in poor health and could not participate. Eighty-six percent of the remaining 839 participants ($n = 719$) were followed. Twenty participants were removed from this sample because they did not participate at the 10-year follow-up. Thus, the total sample size at the 20-year follow-up was 699.

Measures

Depression—Depressive symptoms were measured at each assessment using 18 items from the Health and Daily Living (HDL) scale (Moos, Cronkite, & Finney, 1992) that were derived from the Research Diagnostic Criteria (Spitzer, Endicott, & Robins, 1978). Respondents indicated the frequency with which they experienced each depressive symptom over the past month on a 5-point scale, ranging from 0 (*never*) to 4 (*very often*). The first 10 items of this measure closely parallel the symptoms included in the DSM-IV diagnostic criteria for a major depressive episode. However, these 18 items tap into a broader array of symptoms of depressed mood and ideation (e.g., feeling guilty, worthless, or down on yourself; feeling negative or pessimistic) as well as other depressive features (e.g., crying, feeling resentful, irritable, angry). These HDL items have convergent validity with the Beck Depression Inventory across two time points ($r = .88$ and $.92$; Billings & Moos, 1985). In the present sample, items demonstrated high internal consistency at baseline and the 10- and 20-year follow-ups ($\alpha = .93$ at all three waves).

Criterion Variables—(1) *Medical Conditions* were assessed by a self-report measure derived from the Life Stressors and Social Resources Inventory (LISRES; Moos & Moos, 1994; Moos, 2002). This measure represented the total count of 13 serious medical conditions diagnosed by a physician (e.g., arthritis, cancer, diabetes, high blood pressure, stroke, or ulcer) that participants reported experiencing during the past year. (2) *Drinking Problems* in the last year were assessed by summing affirmative responses to the 17-item Drinking Problems Index (DPI; $\alpha = .80$ to $.88$; Finney, Moos, & Brennan, 1991). DPI items were specifically designed to tap into drinking problems among older adults and gauged negative physical, psychological, and social consequences associated with alcohol use, such as falling, feeling confused, and complaints from family and friends about one's drinking. (3) *Antidepressant Usage* was measured as a dichotomous variable using a single item that inquired about whether or not antidepressants were used frequently in the last year. (4) *Professional Help Seeking* was a dichotomous variable, assessing whether or not a participant had visited a psychiatrist/psychologist or a marriage/family counselor in the past year.

Taxometric Methodology

Taxometric methodology (Meehl, 1995; Waller & Meehl, 1998) was used to examine the underlying structure of late-life depressive symptoms. Rather than relying on tests of statistical significance, taxometric methodology uses consistency tests, which involves analyzing one's data in as many non-redundant ways as possible to see if coherent results are obtained. The goal is to obtain results that either consistently support a categorical structural model—referred to as *taxonic* in the taxometric literature, as the two groups are the *taxon* (e.g., depressed individuals) and *complement* (e.g., non-depressed individuals)—or consistently support a dimensional structural model. Inconsistent results suggest withholding judgment about which of these two models better fits the data (Meehl, 2004).

Selection of Indicators

Taxometric analyses require the presence of multiple non-redundant indicators of the construct of interest. Indicators of depression were created using the 18 HDL depression items. Because single items did not provide enough variability to serve as indicators, groups of items with similar content were summed together. These groups of summed items were formed by first conducting principal axis factoring with oblique rotation. In order to arrive at the best possible set of indicators across the three time points, the items were summed across the three waves of data and then divided by the number of valid data points (e.g., divided by 3 if information was provided at all time points).¹ The liberal Jolliffe criterion (eigenvalue $\geq .7$) was used in

order to retain a sufficient number of factors that could then be used to create the best possible set of non-redundant indicators for a taxometric analysis (Jolliffe, 1972).

This analysis yielded five factors. The items with the highest three factor loadings were summed to create an indicator, with the exception of one factor that only had two items with high factor loadings. These indicators tapped into a broad range of depressive symptoms, including (1) *negative thinking* (e.g., pessimistic/unpleasant thoughts; thoughts about death/suicide), (2) *lethargy and somatic complaints* (e.g., feeling slowed down or fatigued; physical complaints), (3) *low self-esteem and lack of confidence* (e.g., feeling guilty/worthless or inadequate; trouble making decisions), (4) *changes in eating and sleeping habits* (e.g., trouble sleeping or sleeping too much; poor appetite), and (5) *sadness and dejection* (e.g., crying; needing reassurance; feeling sorry for oneself). Four HDL depression items were not used in the creation of these indicators. These items did not clearly load on any one factor, likely because of the broad nature of their content (e.g., angry, feeling blue) and overlap with general distress or anxiety (e.g., psychomotor agitation, irritability).

The five indicators tended to be positively skewed at the baseline (skewness = .58 to 1.0, SE = .07), 10-year follow-up (skewness = .43 to 1.12, SE = .07), and 20-year follow-up (skewness = .30 to 1.13, SE = .09) assessments.

Analytic Plan

General Description—We used the maximum eigenvalue (MAXEIG) procedure to analyze these data (Waller & Meehl, 1998). A schematic representation of the MAXEIG procedure is presented in Figure 1, which shows a hypothetical curve for a taxonic construct and describes key elements of this analysis. As can be seen in this figure, this analysis involves dividing participants into successive sub-samples or *windows* (that can be adjacent or overlapping) along an *input* indicator on the x axis of a graph. The degree of covariation among two or more *output* indicators for each of these consecutive windows is then calculated as an eigenvalue (the multivariate analog of covariance) and plotted on the y-axis for each window. If a construct is taxonic, one window will presumably include a heterogeneous subsample of participants that consists of roughly equal numbers of taxon and complement members, resulting in maximum covariation among the output indicators, which is observed as a peak.

To illustrate why such a pattern would be expected for a taxonic construct, Ruscio and colleagues (2006, p. 125) have suggested considering a hypothetical example in which we wish to test the latent structure of biological sex, using height, voice pitch, and nonverbal sensitivity as our indicators. Within homogeneous subsamples (or windows) made up exclusively of women, we would expect covariation among these indicators to be negligible (e.g., a woman's height would likely be a poor predictor of her voice pitch or sensitivity to nonverbal cues). The same would be true for subsamples made up exclusively of men. However, in a mixed-sex subsample, women would tend to score predictably higher/lower on these indicators than men (e.g., higher pitched voices, shorter in height), resulting in a substantial correlation among the indicators. This same principle can be applied to a construct like late-life depression. However, in the absence of an infallible criterion for establishing group membership (i.e., depressed vs. non-depressed individuals), we must rely on an imperfect indicator (i.e., an input indicator) to sort participants into relatively pure groups of depressed and non-depressed individuals, which can then be used to test for differences in covariance (i.e., among the output indicators) across successive subsamples.

¹Broadly similar factor structures were observed at the three time points when HDL depression items were analyzed separately for each wave.

When latent structure is dimensional, though, there is no reason to expect that covariation among output indicators would differ in any predictable way across ordered subsamples. Instead, we would anticipate that indicators would covary at similar levels across successive subsamples, resulting in a MAXEIG curve that resembles a straight line without a distinct peak.

In all MAXEIG analyses in the present study, one indicator at a time served as the input while the remaining indicators were used as the output in an alternating fashion (such that each indicator served as the input one time), yielding five curves when all five indicators were used. Indicators were standardized, and windows overlapped by 90%. To reduce sampling error resulting from the arbitrary ordering of equal scoring cases, participants were resorted along the input indicator 20 times in each analysis, and the averaged results across these replications were plotted on the y-axis (Ruscio et al., 2006). All analyses were performed using John Ruscio's (2008) taxometric programs. In performing these analyses, we used two different strategies: (1) the inchworm consistency test (ICT; Waller & Meehl, 1998) and (2) a test involving comparisons between results derived from data collected in this study and results derived from taxonic and dimensional comparison data with known latent structures (Ruscio et al., 2006).

Inchworm Consistency Test—The ICT enabled us to directly address two methodological challenges often encountered with community samples: a relatively small base rate of depression and positively skewed indicators. A latent dimension measured with positively skewed indicators can often give rise to results that look very similar to those that might be observed for a small latent taxon. MAXEIG and ICT have been shown to be among the most powerful combinations of analytic strategies when data are characterized by a low base rate and skewed indicators and can often distinguish latent structures more sensitively than other taxometric procedures, as shown in simulation studies (Ruscio & Marcus, 2007; Ruscio & Ruscio, 2004) as well as in the analysis of collected data (Ruscio, Ruscio, & Keane, 2004; Solomon et al., 2006). The ICT involves performing a MAXEIG analysis multiple times with increasing numbers of windows. This consistency test is based on the logic that a true latent taxon will yield more defined peaks (i.e., flattening or decrease in slope after a peak is reached) with increasing numbers of windows; whereas, “pseudo-taxonic curves” (i.e., peaked curves due to indicator skew or other peculiarities of the data) will either appear unchanged or continue to rise as the number of windows increases (Waller & Meehl, 1998). These pseudo-taxonic curves often show up as upward pointing lines without a well-defined peak or decrease in slope (frequently referred to as cusped peaks). In the present study, baseline and 10-year follow-up data were submitted to MAXEIG analyses with 400, 800, and 1200 windows. To approximate the same number of participants per window, 200, 400, and 600 windows were used for the 20-year follow-up data.

Comparison Data—Comparison data were also used as an interpretive aid when making decisions about whether or not a MAXEIG curve appeared dimensional or taxonic. Comparison data sets were created using a bootstrap procedure to mimic many of the parameters of the actual data (e.g., sample size, number and distributions of the indicators, and inter-correlations of the indicators). For the indicators at each of the three time points, two types of comparison data sets were created—one with a clearly dimensional structure and another with an unambiguous taxonic structure that reproduced indicator correlations and distributions within the putative taxon and complement (see Ruscio et al., 2006, for additional information on comparison data).

MAXEIG curves generated from the comparison data were used as a point of comparison when interpreting the shape of the curves generated from the research data (i.e., the actual data collected for this study).² This strategy allowed us to account for peculiarities in the data and reduced the chance of drawing faulty conclusions when the shape of the curve might be

influenced by factors other than the taxonic or dimensional structure of the data, such as indicator skew. For each analysis, 10 comparison data sets were created for both latent structures in order to provide sampling distributions of results for comparison. Beyond visually inspecting the curves, an objective measure of fit termed the comparison curve fit index (CCFI; Ruscio et al., 2006, Formula 7.4, p. 188) was also calculated. This index takes into account the residuals between the plotted points produced by the research data and those produced by the taxonic and dimensional comparison data. This information is then combined to form the CCFI, which yields values that range from 0 to 1. Ruscio and Walters (in press) advise using dual thresholds of .45 and .55 to interpret taxometric results, whereby CCFI values $> .55$ are taken as evidence of a taxonic structure and CCFI values $< .45$ are taken as evidence of a dimensional structure. If results are found to be ambiguous ($.45 \leq \text{CCFI} \leq .55$), judgment is typically withheld. In a recent simulation study that used MAXEIG, this rule of thumb was found to correctly classify 96.1% of samples when CCFI values were outside the ambiguous range (Ruscio & Walters, in press).

Given the widespread use of the base rate consistency test, the mean and standard deviation of estimated base rates of the taxon are also reported to aid data interpretation. In MAXEIG analyses, base rates were estimated by calculating the proportion of individuals beyond the midpoint of the window with the largest eigenvalue (Waller & Meehl, 1998). The base rate consistency test is based on the notion that more consistent base rate estimates imply that a latent taxon is being reliably detected across analyses. Schmidt, Kotov, and Joiner (2004) have suggested that a standard deviation of .10 or less be considered evidence in support of a taxonic structure. The CCFI was given more weight than other numerical indices, as recent simulation studies suggest that other criteria, such as base rate consistency test, are substantially less effective than the CCFI (Ruscio, 2007; Ruscio, Ruscio, & Meron, 2007).

Suitability of the Data—In order to provide a fair test of latent structure, a sample must include a reasonably large number of individuals who are likely members of the putative taxon—in this case depressed individuals. To establish an initial estimate of the base rate of depression, individuals were classified as depressed or non-depressed using the HDL depressive symptoms items that mapped onto the DSM-IV symptom criteria for a major depressive episode (American Psychiatric Association, 2000). Specifically, a participant had to endorse *fairly often* or *often* experiencing five or more of the nine symptoms included in the DSM-IV criteria, and one symptom had to be depressed/sad mood or loss of interest/pleasure in one's usual activities. Using these DSM-IV-based criteria, the base rate of depression was estimated at 7.0% at baseline, 8.5% at the 10-year follow-up, and 8.0% at the 20-year follow-up.

Though the base rate estimates in this study are somewhat smaller than the 10% minimum recommended by Meehl (1995), it has been suggested that greater attention should be placed on the raw number of participants in the putative taxon, as the addition of more complement members in a taxometric analysis has little impact on its ability to distinguish latent structure (Ruscio & Ruscio, 2004). Because our overall sample was much larger than Meehl's (1995) recommended minimum total sample size of 300 participants, we considered the data suitable for analysis.

²Simulated data were also used to examine the ability of the chosen indicators to correctly classify data sets that mimicked the properties of the research data but had known latent structures. These preliminary analyses revealed that taxometric procedures, such as Latent Mode (L Mode) and mean above and mean below a cut (MAMBAC), were not powerful enough to clearly distinguish simulated taxonic and dimensional data, given the unique properties of the research data (e.g., small base rate, positively skewed indicators). Only the MAXEIG procedure was able to correctly identify the known latent structures in the taxonic and dimensional comparison data, providing additional support for the validity of the chosen indicators and plan of analysis.

To be considered valid for a MAXEIG analysis, indicators must also separate the individuals in the proposed groups with a large effect size. In addition, the inter-correlations among indicators within the hypothesized taxon (depressed individuals) and complement (non-depressed individuals) must be substantially lower than the inter-correlations within the entire sample (Meehl, 1995, Ruscio et al., 2006). This requirement stems from the fact that a taxonic peak in the curve can be obscured if the covariation among the indicators is high within fairly homogeneous groups of participants (perhaps due to indicator redundancy or shared method variance), which can sometimes result in pseudo-dimensional curves (i.e., a taxonic construct with a flattened MAXEIG curve that appears dimensional).

Using the depressed and non-depressed groups established with the DSM-IV-based criteria, the five indicators used in this study were found to separate depressed from non-depressed individuals by $d = 1.95$ to 2.01 on average across the three waves of data. Inspection of the inter-correlations among indicators within the depressed and non-depressed groups revealed that indicator 3 (low self-esteem and lack of confidence) and indicator 5 (sadness and dejection) were highly correlated with at least two other indicators within groups with correlations as high as $.54$ in the depressed group and $.68$ in the non-depressed group across the three waves of data collection. Among the remaining three indicators (indicators 1, 2, and 4), inter-correlations within the depressed (Mean $r = .08-.18$) and non-depressed (Mean $r = .40-.43$) groups were substantially lower than the inter-correlations within the entire sample (Mean $r = .51-.53$). Hence, these three strongest indicators (indicators 1, 2, and 4) were analyzed separately. Although this restricted set of indicators was advantageous from a statistical standpoint, it was perhaps less desirable from a conceptual standpoint, in that these three indicators likely did not tap into the construct of interest as fully as the complete set of five indicators. In order to address this issue and to provide an additional consistency test, analyses were performed using all five indicators (referred to as the 5 indicators) as well as with the restricted set of three indicators (referred to as the 3 best indicators).

Results

Inchworm Consistency Test

ICT curves are presented in Figures 2 and 3. Within each graph, solid lines represent the average of all MAXEIG curves, and the individual curves are presented as broken lines. As shown in Figure 2, when the 5 indicators were used, MAXEIG analyses tended to yield slight right-ended peaks for averaged curves at all three time points; however, the peaks at the 10-year follow-up were less visible. However, increasing the number of windows did not appear to yield more distinct peaks for analyses with these 5 indicators.

When MAXEIG analyses were performed with the 3 best indicators, peaks tended to be much more prominent across all three waves of data (see Figure 3). At baseline and the 20-year follow-up, MAXEIG analyses with 400 and 200 windows, respectively, yielded curves with cusped right-ended peaks, but each of these peaks showed a visible decrease in slope with increasing numbers of windows, which supports a taxonic conceptualization. For the 10-year follow-up data, the right-ended peak at 400 windows was already somewhat flattened and then reached a more prominent point at 800 and 1200 windows.

Results for Comparison Data

The shape of these MAXEIG curves were compared to curves derived from comparison data sets with either a taxonic or dimensional structure. Given that there was some evidence of a latent taxon in the ICT analyses, cases were first reclassified as probable taxon and complement members at each wave using results from an initial MAXEIG analysis in order to better approximate the base rate of a potential taxon in the current sample. Specifically, Bayes'

theorem was used to compute the probability of taxon membership for each participant, and individuals were assigned to the more probable group (see Ruscio et al., 2006, pp. 323–325, for a detailed description of this procedure). The percentage of individuals classified as probable taxon members was 1.6%, 6.3%, and 3.9% for the 5 indicators and 2.5%, 5.4%, and 4.6% for the 3 best indicators at baseline, 10-year follow-up and 20-year follow-up, respectively. This classification method was used when generating taxonic comparison data that attempted to reproduce indicator correlations and distributions within the putative taxon and complement. In these analyses, 800 windows were used for baseline and 10-year follow-up data, and 400 windows were used for the 20-year follow-up data, in order to provide a maximally sensitive analysis without sacrificing the stability of the curves by having too few participants in each window.

Five Indicators—Results for the five indicators are presented in Figure 4. At baseline, all five of the MAXEIG curves generated from these indicators yielded a right-ended peak, and the average curve representing these results appeared to map onto the curves generated from the taxonic comparison data better than the curves generated from the dimensional comparison data—an observation supported by a CCFI value of .62. Base rate estimates were also highly stable across the five MAXEIG curves ($M = .08$, $SD < .01$), further supporting a taxonic solution.

For the five indicators at the 10-year follow-up, three of the five curves yielded visible right-ended peaks, and when compared to the comparison data, the averaged curve of the research data appeared to map onto the curves generated from the taxonic comparison data, which was supported by a CCFI of .58. Consistent with a taxonic structure, base rate estimates were also stable across the five curves at the 10-year follow-up ($M = .13$, $SD = .05$).

At the 20-year follow-up assessment, each of the five curves exhibited a right-ended peak. Although the differences between the curves generated from the taxonic and dimensional comparison data were not particularly striking at this time point, the average curve of the research data more closely resembled the taxonic comparison curve, and a CCFI of .55 was obtained. Base rate estimates were also highly stable for the 20-year follow-up data ($M = .13$, $SD = .02$), providing additional support for a taxonic structure.

Three Best Indicators—As shown in Figure 5, peaks were generally more prominent when the 3 best indicators were used. At baseline, all three MAXEIG curves yielded right-ended peaks, and the averaged curve more closely resembled the curve generated from the taxonic comparison data compared to the curve from the dimensional comparison data, as indicated by a CCFI of .60. Also consistent with a taxonic structure, the base rate estimates were found to be highly stable across the three MAXEIG curves ($M = .07$, $SD = .02$).

For the 10-year follow-up data, all three of the MAXEIG curves produced right-ended peaks, and the averaged curve yielded a prominent peak that generally resembled the curve derived from the taxonic comparison data (CCFI = .58). The base rate estimates were also fairly stable at this time point ($M = .11$, $SD = .06$).

The 20-year follow-up data yielded right-ended peaks across all three curves, and the averaged curve appeared to approach the crest of a peak. However, differences between the curves generated from the taxonic and dimensional comparison data sets were slight and did not offer much clarity as an interpretive aid. Despite this ambiguity, a CCFI value of .55 was obtained, indicating that the curve of the research data at the 20-year follow-up resembled the curve of the taxonic comparison data somewhat more than the curve from the dimensional comparison data. The three MAXEIG curves also generated stable base rate estimates at this time point ($M = .11$, $SD = .03$).

Follow-up Tests

Because taxometric analyses revealed evidence supporting the existence of a possible latent taxon, follow-up analyses were performed that examined which symptoms of a major depressive episode (represented by the first 10 items of the HDL depression measure) best characterized individuals who had been classified as likely taxon members (established using Bayes' theorem). Those classified as likely taxon members in the MAXEIG analyses involving the full set of 5 indicators were used in these analyses, given that this indicator set captured a broader range of depressive symptoms.

Although those assigned as probable taxon members were likely to also have been diagnosed as depressed using DSM-IV-based criteria, there was not perfect overlap between these two groups, as shown in Table 1. Notably, the groups of likely taxon members were substantially smaller than the group of participants initially classified as depressed using DSM-IV-based criteria, which could indicate that the identified taxon represents only a subgroup of depressed individuals. If the probable taxon members do in fact make up a distinct subgroup, Table 1 would suggest that DSM-IV-based depressed older adults (as might be encountered in clinical practice) are heterogeneous, with 21.1% to 50.9% exhibiting this taxonic form of depression and 49.1% to 78.9% not being recognized as taxon members. It should also be noted that the ratio of probable taxon members to DSM-IV-based depressed individuals appeared to differ across the three waves of data collection. At baseline, in particular, this ratio was weighted toward having a lower proportion of taxon members, which could suggest that DSM-IV-based depression becomes more characterized by this taxonic form after age 65.

To determine how taxon members might differ from non-depressed individuals and from individuals likely to meet DSM criteria for a major depressive episode, multinomial regression analyses were performed. For these analyses, the outcome variable was divided into three categories: (1) the *taxon members* (those assigned as probable taxon members using Bayes' theorem), (2) a *DSM-only depressed group* (those who were initially classified as depressed using DSM-IV-based criteria but were not categorized as likely taxon members), and (3) a *non-depressed group* (those who were categorized as non-depressed by the DSM-IV-based criteria and categorized as complement members using Bayes' theorem).

Multinomial logistic regression analyses were run for all three waves of data (see Table 2). In these analyses, the 10 HDL items representing the symptoms of a major depressive episode were used as the independent variables. We also included several possible confounding variables in our models, including age (in years), sex (1 = male, 2 = female), ethnicity (1 = Caucasian, 2 = ethnic minority), years of education, and number of medical conditions. Baseline values for age, sex, ethnicity, and years of education were used in all analyses because their variability was, for the most part, fixed across the three waves of data collection. In contrast, unique values at each time point were used for medical conditions and the 10 HDL depression items. Taxon members were used as the reference category for the outcome variable.

For comparisons made between the non-depressed group and taxon members, tiredness/fatigue and loss of interest in usual activities consistently predicted group membership at each of the three time points, with taxon members being more likely to experience these symptoms. For comparisons made between the DSM-only depressed group and taxon members, only thoughts of death/suicide were able to predict taxon membership across all three waves of data, with taxon members being more likely to have these types of thoughts. At the 20-year follow-up, men were more likely to be taxon members (when contrasted against the non-depressed group), as were individuals with more years of education (when contrasted against both the non-depressed and DSM-only groups).

Criterion Variable Analyses

A secondary research question concerned the relative utility of treating depression as a continuous variable (represented as a total HDL depression scale score) or as a dichotomous variable in predicting several relevant criterion variables. Since we found evidence supporting the existence of a possible latent taxon, we used the observed boundary between taxon and complement members as the basis for creating our dichotomous variable. In particular, the dichotomous depression variable was represented by “taxon membership” (i.e., individuals classified as probable taxon or complement members) as established by Bayes’ theorem in the MAXEIG analyses involving the 5 indicators. In order to objectively compare the magnitude of the correlations with the criterion variables for the dichotomized taxon membership variable versus those derived from the continuous total 18-item HDL depression variable, we used Williams’ (1959) $t(t_w)$ test. Williams’ t can be used to test for differences between non-independent correlations derived from a common third variable (i.e., the criterion variables) and the same sample. This statistic can be interpreted the same as any t test with $N - 3$ degrees of freedom.

Correlations are shown in Table 3. In all but five instances, Williams’ t tests were statistically significant and indicated that HDL depression scores were more strongly associated with the criterion variables in the concurrent and predictive analyses, compared to correlations based on the dichotomized taxon membership variable. Three of the five nonsignificant Williams’ t tests were obtained for correlations between the depression measures and 20-year drinking problems. However, in these cases, the similarity in the performance of these variables was not particularly surprising, given that neither of the depression variables (at baseline, 10-years, and 20-years) were significantly associated with problem drinking at the 20-year follow-up—likely due to the infrequency and limited variability in drinking problems at this time point. In addition, both the dichotomized and continuous depression variables at baseline and the 10-year follow-up significantly predicted help-seeking at the 20-year follow-up. However, neither continuous nor dichotomized depression variables performed significantly better than the other in predicting 20-year help-seeking at baseline ($t_w(696) = .87, p = .19$) or at the 10-year follow-up ($t_w(696) = -.03, p = .49$).

Discussion

Overall, our findings favored a taxonic (categorical) conceptualization of late-life depressive symptoms. Peaked MAXEIG curves were observed at all three waves. These curves also consistently generated stable base rate estimates and generally mapped onto the curves derived from the taxonic comparison data better than those from the dimensional comparison data. Nevertheless, in some cases results were not striking (e.g., CCFI values were somewhat ambiguous at the 20-year follow-up), indicating that these findings should be interpreted cautiously until they can be replicated with other older adult samples. In addition, the inchworm consistency tests for the 5 indicators appeared less peaked, compared to those observed for the 3 best indicators, possibly indicating that some indicators were better than others at identifying the taxon. Indeed, when probable taxon members were examined in follow-up analyses, they were not necessarily characterized by the full range of DSM-IV criteria for a major depressive episode. For example, compared to those classified as depressed using DSM-IV-based criteria, taxon members were primarily characterized by thoughts about death and suicide. Other potential markers of this group included loss of interest in usual activities and fatigue/tiredness. However, this group was not particularly characterized by sadness. This finding fits with past research suggesting that older adults may be less likely than other age groups to report affective symptoms of depression (Gallo et al., 1994). It may also help explain results favoring a categorical conceptualization of late-life depressive symptoms, given that past research has found affective symptoms of depression to be distributed more as a dimension and somatic

complaints (e.g., tiredness/fatigue) to be distributed more categorically (Beach & Amir, 2003).

Beyond these findings, it is notable that the number of participants identified as probable taxon members was substantially less than the number of participants identified as depressed using DSM-IV-based criteria. Thus, it could be that the identified taxon represents a specific subset of depressed older adults. In some ways, the symptom profile of this group maps onto the symptoms of the melancholic subtype of depression (e.g., loss of interest, fatigue), and several previous studies have supported a taxonic structure for this subtype (Haslam, 2003). However, taxon members were primarily characterized by suicidal/death ideation, which does not necessarily exemplify melancholic depression (Rush & Weissenburger, 1994). Thus, this group could represent the proposed hopelessness subtype of depression (Abramson, Alloy, & Metalsky, 1989), which is often characterized by suicidal ideation along with other symptoms such as psychomotor retardation, fatigue, and a negative outlook (Joiner et al., 2001). However, two taxometric studies have supported a dimensional conceptualization of hopelessness depression (Haslam & Beck, 1994; Whisman & Pinto, 1997), which raises the possibility that the symptoms of this subtype could be distributed differently among older adults.

Despite the support found for a taxonic structure, variability within the taxon and complement appeared to be meaningful. Indeed, compared to the dichotomized taxon membership variable, continuous total HDL depression scores were more strongly associated with most of the criterion variables. Thus, although there may be a dividing line that separates older depressed from older non-depressed individuals, it still makes sense to treat late-life depressive symptoms continuously in many situations, particularly when predicting outcomes like problem drinking or medical conditions. However, it is notable that the dichotomized taxon membership variable (at baseline and at the 10-year follow-up) performed similarly as the continuous depression measure when predicting professional help seeking at the 20-year follow-up. This finding suggests that there may be some clinically relevant outcomes for which a continuous depression measure explains little additional variance, beyond that explained by an appropriately dichotomized depression variable.

Limitations, Implications, and Future Directions

The present sample was comprised primarily of Caucasian individuals and had an upper age limit of 85 years at the 20-year follow-up assessment. Further, this sample's parent study was designed to focus on late-life drinking behavior. Accordingly, lifetime non-drinkers were excluded from the sample at study recruitment, and the alcohol consumption of the recruited sample may not have been representative of the population of late-middle-aged adults. There are divergent findings regarding the relationship between alcohol consumption and depressive symptoms (e.g., Schutte et al., 1995; Wilsnack, 1991), but the recruitment procedures used in the parent study may have yielded a sample that was not representative of the overall population of older adults with respect to depressive symptoms. Future research should ascertain the generalizability of the present findings to samples of adults whose ages extend beyond 85 years, to more ethnically diverse samples, and to samples whose range and distribution of depressive symptoms are known to closely resemble those in the overall population of older adults.

In addition, some of the findings in this study indicated that taxon members at the 20-year follow-up might be more likely to be men (see Table 2). This finding, along with our discovery that thoughts of suicide and death characterize the taxon, could reflect a tendency for older men to have more thoughts than women about suicide/death (Osgood, 1992). We also found that taxon members at the 20-year follow-up had more formal education than non-depressed participants and participants meeting criteria for a DSM-IV-based definition of depression (but not identified as a taxon member). Future research with other cohorts might usefully explore if education is a consistent predictor of taxon membership. It is also noteworthy that the present

sample was comprised of community residents, who were not particularly at-risk for experiencing depressive symptoms. This type of sample has the advantage of representing a wide spectrum of variability in symptoms but often yields a smaller sized taxon that can be harder to identify. In the present investigation, comparison data sets were used as interpretive aids to determine whether the shapes of the curves were due to latent structure or unrelated peculiarities of the data. However, future studies that recruit samples with larger base rates of depressed individuals may yield more striking results.

This study was also limited by the fact that only a single self-report measure of depression was used to create indicators of depressive symptoms in the taxometric analyses. Future studies of the underlying structure of late-life depressive symptoms would do well to test a broader range of symptoms and perhaps employ multiple methods of data collection (e.g., clinician administered measures, reports of family/friends), which could reduce “statistical noise” among indicators due to shared methods. The use of multiple measures could also help clarify which symptoms characterize a taxonic form of late-life depression. For example, studies that employ multiple measures of depression might establish taxon membership with one measure and then examine which symptoms typify taxon membership using a different one.

It should also be noted that our item that assessed thoughts about suicide and/or death, which seemed to characterize probable taxon members, did not make distinctions between suicidal ideation, death ideation, or benign thoughts about death. Though some studies have found death ideation and suicidal ideation to be similarly associated with psychological and health problems among older adults (Heisel & Flett, 2006; Scocco et al., 2001), other studies have reported unique associations between death ideation and medical comorbidities (Bartels et al., 2002; Kim, Bogner, Brown, & Gallo, 2006), raising the possibility that the identified taxon might represent a medically ill group. The fact that number of medical conditions was not consistently associated with taxon membership would suggest that these individuals were not characterized by physical illnesses; however, we could not definitively rule out this possibility.

On a similar note, individuals in this study were not excluded due to the presence of a medical condition, recent bereavement, or substance use problems. As an older adult sample, most participants endorsed at least one of these problems, highlighting the difficulty of isolating a single etiological factor. Therefore, we were not able to differentiate between a major depressive episode and other mood disorders, such as substance-induced mood disorder or mood disorder due to a general medical condition. As a result, this study may be considered an examination of the latent structure of late-life depressive symptoms generally rather than the symptoms associated with any one specific diagnostic category.

We also acknowledge that the question of whether to treat a construct as continuous or categorical is not entirely an empirical issue. Indeed, there are often theoretical or practical concerns (e.g., issues related to insurance/billing) that influence a researcher or clinician’s decision about how to conceptualize late-life depressive symptoms. Notwithstanding these issues, this study is based on the premise that scientific methods can be used to objectively observe the underlying structure of a construct, and these observations provide additional information that can be used to make decisions about measurement and assessment.

In particular, if future studies are able to identify a taxonic boundary similar to the one observed in this study, this dividing line could serve as basis for defining clinically meaningful change in practice and research settings. Further explication of the late-life depressive symptoms that best characterize this group could also provide supplementary diagnostic guidelines, which could complement existing DSM criteria and help identify these individuals. Indeed, early identification could be important in these cases. For example, in this study a dichotomous variable (based on the identified taxon) predicted professional help seeking at the 20-year

follow-up about as well as a continuous depression measure, indicating that the individuals who made up this group might have been particularly at risk for long-term problems that led them to seek professional help or be referred for help. Additionally, given that the identified taxon members were primarily characterized by thoughts of death/suicide, it seems possible that this group may be at heightened risk for self-harm. Considering that older adults are at greater risk for suicide than other age groups (National Institute of Mental Health, 2007), clinicians would do well to be watchful for individuals exhibiting symptom profiles similar to these taxon members.

Of course, these findings also suggest that a continuous conceptualization might still provide important predictive advantages in many circumstances, especially for measuring sub-threshold late-life depressive symptoms in certain populations such as those with medical conditions or drinking problems. Future studies would do well to further examine the relative merits of conceptualizing late-life depressive symptoms dichotomously vs. continuously when predicting a variety of other clinically relevant outcomes. For instance, it seems possible that an appropriately dichotomized depression variable might predict outcomes, like hopelessness or the will to live, more cleanly than a continuous measure, particularly given that the identified taxon in this investigation was largely characterized by thoughts of death/suicide. Future studies might also examine the implications of a categorical conceptualization for treatment, as it is possible that identified taxon members could respond better to particular interventions, compared to sub-threshold cases of depression.

Acknowledgments

Preparation of this article was supported by NIAAA Grant AA15685 and by the Department of Veterans Affairs Health Services Research and Development Service and Office of Academic Affiliation. The views expressed here are the authors' own and do not necessarily represent the views of the Department of Veterans Affairs or the United States Government. We would like to thank John Ruscio for offering comments on an earlier draft of this paper and s Bernice Moos for her assistance in data collection and preparation.

References

- Abramson LY, Metalsky GI, Alloy LB. Hopelessness depression: A theory based subtype of depression. *Psychological Review* 1989;96:358–358.
- Alexopoulos G, Meyers B, Young R, Campbell S, Silbersweig D, Charlson M. The “vascular depression” hypothesis. *Archives of General Psychiatry* 1997;54:915–922. [PubMed: 9337771]
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4. Washington, DC: Author; 2000. text revision
- Atkinson R. Depression, alcoholism and ageing: A brief review. *International Journal of Geriatric Psychiatry* 1999;14:905–910. [PubMed: 10556860]
- Bartels SJ, Coakley E, Oxman TE, Constantino G, Oslin D, Chen H, et al. Suicidal and death ideation in older primary care patients with depression, anxiety, and at-risk alcohol use. *American Journal of Geriatric Psychiatry* 2002;10:417–427. [PubMed: 12095901]
- Beach SRH, Amir N. Is depression taxonic, dimensional, or both? *Journal of Abnormal Psychology* 2003;112(2):228–236. [PubMed: 12784832]
- Beach SRH, Amir N. Self-reported depression is taxonic. *Journal of Psychopathology and Behavioral Assessment* 2006;28:171–178.
- Billings AG, Moos RH. Life stressors and social resources affect posttreatment outcomes among depressed patients. *Journal of Abnormal Psychology* 1985;94:140–153. [PubMed: 3998282]
- Brennan PL, Moos RH. Life stressors, social resources, and late-life problem drinking. *Psychology and Aging* 1990;5:491–501. [PubMed: 2278671]
- Finney JW, Moos RH, Brennan PL. The Drinking Problems Index: A measure to assess alcohol-related problems among older adults. *Journal of Substance Abuse* 1991;3:395–404. [PubMed: 1821293]

- Gallo JJ, Anthony JC, Muthén BO. Age differences in the symptoms of depression: A latent trait analysis. *Journals of Gerontology* 1994;49:P251–P264. [PubMed: 7963280]
- Gallo JJ, Rabins PV, Anthony JC. Sadness in older persons: 13-year follow-up of a community sample in Baltimore, Maryland. *Psychological Medicine* 1999;29(2):341–350. [PubMed: 10218925]
- Gilbert, P. Varieties of submissive behavior as forms of social defense: Their evolution and role in depression. In: Sloman, L.; Gilbert, P., editors. *Subordination and defeat: An evolutionary approach to mood disorders and their therapy*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 2000. p. 3–45.
- Gottfries CG. Is there a difference between elderly and younger patients with regard to the symptomatology and aetiology of depression. *International Clinical Psychopharmacology* 1998;13:S13–S18. [PubMed: 9817615]
- Grunebaum MF, Oquendo MA, Manly JJ. Depressive symptoms and antidepressant use in a random community sample of ethnically diverse, urban elder persons. *Journal of Affective Disorders* 2008;105:273–277. [PubMed: 17532052]
- Haslam N. Categorical versus dimensional models of mental disorder: The taxometric evidence. *Australian and New Zealand Journal of Psychiatry* 2003;37:696–704. [PubMed: 14636384]
- Haslam N, Beck AT. Subtyping major depression: A taxometric analysis. *Journal of Abnormal Psychology* 1994;103:686–686. [PubMed: 7822569]
- Heisel MJ, Flett GL. The development and initial validation of the Geriatric Suicide Ideation Scale. *American Journal of Geriatric Psychiatry* 2006;14:742–751. [PubMed: 16943171]
- Joiner TE, Steer RA, Abramson LY, Alloy LB, Metalsky GI, Schmidt NB. Hopelessness depression as a distinct dimension of depressive symptoms among clinical and non-clinical samples. *Behaviour Research and Therapy* 2001;39:523–523. [PubMed: 11341249]
- Jolliffe I. Discarding variables in principal component analysis. *Applied Statistics* 1972;21:160–173.
- Kim YA, Bogner HR, Brown GK, Gallo JJ. Chronic medical conditions and wishes to die among older primary care patients. *International Journal of Psychiatry in Medicine* 2006;36:183–198. [PubMed: 17154148]
- Meehl PE. Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist* 1995;50:266–275. [PubMed: 7733538]
- Meehl PE. What's in a taxon? *Journal of Abnormal Psychology* 2004;113:39–43. [PubMed: 14992655]
- Moos, R. *The Life Stressors and Social Resources Inventory and the Coping Responses Inventory: An annotated bibliography*. 2. Palo Alto, CA: Department of Veterans Affairs, Center for Health Care Evaluation; 2002.
- Moos, RH.; Cronkite, RC.; Finney, JW. *Health and Daily Living Form manual*. 2. Palo Alto, CA: Consulting Psychologists Press; 1992.
- Moos, R.; Moos, B. *Life Stressors and Social Resources Inventory: Adult Form manual*. Odessa, FL: Psychological Assessment Resources; 1994.
- Moos RH, Schutte KK, Brennan PL, Moos BS. Older adults' consumption and late-life drinking problems: A 20-year perspective. *Addiction* 2009;104:1293–1302. [PubMed: 19438836]
- National Institute of Mental Health. Older adults: Depression and suicide facts. 2007. Retrieved July 15, 2008 from <http://www.nimh.nih.gov/health/publications/older-adults-depression-and-suicide-facts.shtml>
- Osgood, N. *Suicide in later life: Recognizing the warning signs*. New York, NY: Lexington Books; 1992.
- Parker G, Roy K, Hadzi-Pavlovic D, Mitchell P, Wilhelm K. Distinguishing early and late onset non-melancholic unipolar depression. *Journal of Affective Disorders* 2003;74:131–138. [PubMed: 12706514]
- Parslow RA, Jorm AF. Who uses mental health services in Australia? An analysis of data from the National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry* 2000;34:997–1008. [PubMed: 11127632]
- Ruscio J. Taxometric analysis: An empirically grounded approach to implementing the method. *Criminal Justice and Behavior* 2007;34:1588–1622.
- Ruscio, J. *Taxometric programs for the R computing environment: User's manual*. 2008. Retrieved December 16, 2008, from <http://www.taxometricmethod.com>

- Ruscio J, Brown TA, Ruscio AM. A taxometric investigation of DSM-IV major depression in a large outpatient sample. *Assessment*. in press.
- Ruscio, J.; Haslam, N.; Ruscio, AM. *Introduction to the taxometric method: A practical guide*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 2006.
- Ruscio J, Marcus DK. Detecting small taxa using simulated comparison data: A reanalysis of Beach, Amir, and Bau's data. *Psychological Assessment* 2007;19:241–246. [PubMed: 17563206]
- Ruscio J, Ruscio AM. Clarifying boundary issues in psychopathology: The role of taxometrics in a comprehensive program of structural research. *Journal of Abnormal Psychology* 2004;113:24–38. [PubMed: 14992654]
- Ruscio J, Ruscio AM, Keane TM. 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* 2004;113:145–154. [PubMed: 14992667]
- Ruscio J, Ruscio AM, Meron M. Applying the bootstrap to taxometric analysis: Generating empirical sampling distributions to help interpret results. *Multivariate Behavioral Research* 2007;42:349–386.
- Ruscio J, Walters GD. Using comparison data to differentiate categorical and dimensional data by examining factor score distributions: Resolving the mode problem. *Psychological Assessment*. in press.
- Ruscio J, Zimmerman M, McGlinchey JB, Chelminski I, Young D. Diagnosing major depressive disorder XI: A taxometric investigation of the structure underlying DSM-IV symptoms. *Journal of Nervous and Mental Disease* 2007;195:10–19. [PubMed: 17220734]
- Rush AJ, Weissenburger JE. Melancholic symptom features and DSM-IV. *American Journal of Psychiatry* 1994;151:489–489. [PubMed: 8147445]
- Schmidt, NB.; Kotov, R.; Joiner, TE, Jr. *Taxometrics: Toward a new diagnostic scheme for psychopathology*. Washington, DC, US: American Psychological Association; 2004.
- Schutte K, Moos R, Brennan P. Depression and drinking behavior among women and men: A three-wave longitudinal study of older adults. *Journal of Consulting and Clinical Psychology* 1995;63:810–822. [PubMed: 7593875]
- Scocco P, Meneghel G, Caon F, Dello Buono M, De Leo D. Death ideation and its correlates: Survey of an over-65-year-old population. *Journal of Nervous and Mental Disease* 2001;189:210–218. [PubMed: 11339316]
- Seligman M. Comment and integration. *Journal of Abnormal Psychology* 1978;87:165–179. [PubMed: 649850]
- Solomon A, Ruscio J, Seeley JR, Lewinsohn PM. A taxometric investigation of unipolar depression in a large community sample. *Psychological Medicine* 2006;36:973–985. [PubMed: 16700963]
- Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry* 1978;35:773–782. [PubMed: 655775]
- van den Berg M, Oldehinkel A, Bouhuys A, Brilman E, Beekman A, Ormel J. Depression in later life: Three etiologically different subgroups. *Journal of Affective Disorders* 2001;65:19–26. [PubMed: 11426505]
- Waller, NG.; Meehl, PE. *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks, CA, US: Sage Publications, Inc; 1998.
- Whisman MA, Pinto A. Hopelessness depression in depressed inpatient adolescents. *Cognitive Therapy and Research* 1997;21:345–358.
- Williams EJ. The comparison of regression variables. *Journal of the Royal Statistical Society, Series B* 1959;21:396–399.
- Wilsnack SC, Klassen AD, Schur BE, Wilsnack RW. Predicting onset and chronicity of women's problem drinking: A five-year longitudinal analysis. *American Journal of Public Health* 1991;81:305–318. [PubMed: 1994739]
- Zeiss, AM.; Lewinsohn, PM.; Rohde, P. Functional impairment, physical disease, and depression in older adults. In: Kato, PM.; Mann, T., editors. *Handbook of diversity issues in health psychology*. New York, NY, US: Plenum Press; 1996. p. 161-184.

Figure 1.
Schematic diagram illustrating key elements of the MAXEIG procedure for a hypothetical taxonic curve.

Figure 2.

MAXEIG curves for the inchworm consistency test involving the full set of 5 indicators. Solid lines represent the averaged curve, and dotted lines represent the individual curves.

Figure 3.

MAXEIG curves for the inchworm consistency test involving the 3 best indicators. Solid lines represent the averaged curve, and dotted lines represent the individual curves.

Figure 4.

In these panels, averaged curves are presented as a single dark line that summarizes the results obtained using the full set of 5 indicators. This averaged curve is superimposed over the curves obtained using the simulated data. The curves derived from the simulated taxonic and dimensional data sets are presented as more narrow lines. For each simulated latent structure, 10 samples were drawn in order to provide a sampling distribution of results, which are represented as an upper and lower line.

Figure 5.

In these panels, averaged curves are presented as a single dark line that summarizes the results obtained using the 3 best indicators. This averaged curve is superimposed over the curves obtained using the simulated data. The curves derived from the simulated taxonic and dimensional data sets are presented as more narrow lines. For each simulated latent structure, 10 samples were drawn in order to provide a sampling distribution of results, which are represented as an upper and lower line.

Table 1

Convergence between Taxon Membership and DSM-IV-based Diagnoses

		Probable taxon members	Probable complement members	Total	Chi square	Sensitivity/Specificity
Baseline	DSM depressed	19	71	90	$\chi(1) = 229.14$ $p < .001$	Sensitivity: 21.11% Specificity: 99.83%
	DSM non-depressed	2	1197	1199		
	Total	21	1268	1289		
10-year follow-up	DSM depressed	56	54	110	$\chi(1) = 406.67$ $p < .001$	Sensitivity: 50.91% Specificity: 97.88%
	DSM non-depressed	25	1154	1179		
	Total	81	1208	1289		
20-year follow-up	DSM depressed	18	38	56	$\chi(1) = 131.11$ $p < .001$	Sensitivity: 32.14% Specificity: 98.60%
	DSM non-depressed	9	634	643		
	Total	27	672	699		

Note. Probable taxon members were identified using Bayes' theorem and all five indicators. Sensitivity is the percent of DSM-depressed individuals who are taxon members; specificity is the percent of DSM-non-depressed individuals who are not taxon members. Low sensitivity values could indicate that probable taxon members represent a distinct subgroup of depressed individuals.

Table 2
Multinomial Logistic Regression Analyses Predicting Group Membership at Baseline, 10-year Follow-up, and 20-year Follow-up

Symptom	Baseline (N = 1282)				10-year Follow-up (N = 1274)				20-year Follow-up (N = 692)			
	Non-Depressed Group vs. Taxon Members	DSM-Only Depressed Group vs. Taxon Members	Odds Ratio	SE	Non-Depressed Group vs. Taxon Members	DSM-Only Depressed Group vs. Taxon Members	Odds Ratio	SE	Non-Depressed Group vs. Taxon Members	DSM-Only Depressed Group vs. Taxon Members	Odds Ratio	SE
Age (years)	-.03	.13	.98	.12	.04	.07	1.04	.07	-.28	.18	.75	.14
Sex	-.12	.79	.89	.71	-.20	.39	.82	.42	2.58	1.19	13.16*	.91
Ethnicity	.30	.55	1.35	.47	.09	.72	1.09	.71	.40	3.11	1.50	2.87
Years of education	.01	.16	1.01	.13	.08	.08	1.09	.09	-1.20	.38	.30**	.27
Medical conditions	-.01	.25	.99	.21	.05	.10	1.05	.11	-.57	.28	.57*	.27
Feeling sad/blue	-2.17	.60	.11***	.54	-.29	.22	.75	.22	-2.22	.76	.11**	.51
Poor appetite	-.16	.34	.85	.28	-.01	.16	.99	.16	-.37	.48	.69	.36
Sleep problems	-.38	.42	.69	.36	-.19	.18	.83	.22	-1.74	.75	.18*	.41
Tiredness/fatigue	-1.82	.62	.16**	.56	-.97	.27	.38***	.31	-5.65	1.33	.00***	.97
Unable to sit still	-.09	.38	.91	.34	-.21	.19	.81	.19	.23	.51	1.26	.36
Slowed down	-.90	.47	.41	.43	-.60	.21	.55***	.23	-3.54	.96	.03***	.61
Loss of interest	-.88	.39	.41*	.34	-.57	.22	.57**	.24	-2.36	.78	.10**	.46
Feeling guilty/worthless	-.29	.49	.75	.44	-.67	.18	.51***	.18	-1.30	.83	.27	.52
Trouble concentrating	-1.26	.48	.29**	.44	-.97	.22	.38***	.22	-1.08	.57	.34	.47
Thoughts of death/suicide	-2.00	.43	.14***	.39	-.95	.18	.39***	.19	-.89	.53	.41	.43

Note. In each analysis, taxon members were used as the reference group. As a result, odds ratios in this table represent the likelihood of not being a taxon member. Sample sizes reported above are somewhat reduced due to missing data for particular items. Baseline values for age, sex (1 = male, 2 = female), ethnicity (1 = Caucasian, 2 = ethnic minority), and years of education were used in these analyses. In contrast, unique values at each time point were used for medical conditions and the HDL depression items.

*
 $p < .05$,

**
 $p < .01$,

 $p < .001$

Table 3
Correlations for Taxon Membership, Total Depression Scores, and Criterion Variables

Cross-sectional Analyses				
	Medical Problems (MP)	Problem Drinking (PD)	Antidepressant Use (AU)	Professional Help Seeking (PHS)
Mean (SD) or %				
Baseline	1.47(1.46)	1.34(2.75)	8.9%	15.5%
10-yr follow-up	1.94(1.69)	0.73(1.96)	14.8%	11.2%
20-yr follow-up	2.32(1.80)	0.44(1.31)	16.2%	9.2%
Taxon Membership (Baseline)	.08**	.05	.20***	.15***
Total HDL Depression (Baseline)	.26***	.23***	.35***	.29***
Taxon Membership (10-yr follow-up)	.13***	.07*	.21***	.17***
Total HDL Depression (10-yr follow-up)	.30***	.15***	.33***	.30***
Taxon Membership (20-yr follow-up)	.20***	.04	.18***	.09*
Total HDL Depression (20-yr follow-up)	.34***	.06	.35***	.17***

Predictive Analyses							
	10-year Follow-up			20-year Follow-up			
	MP	PD	AU	PHS	MP	PD	AU
Taxon Membership (Baseline)	.14***	.02	.16***	.15***	.14***	-.01	.19***
Total HDL Depression (Baseline)	.27***	.15***	.28***	.21***	.23***	.02	.31***
Taxon Membership (10-yr follow-up)	--	--	--	--	.14***	-.03	.21***
Total HDL Depression (10-yr follow-up)	--	--	--	--	.28***	.02	.32***

Note. N = 1289 for analyses involving baseline and 10-year follow-up data, and N = 699 for analyses involving 20-year follow-up data. Means and standard deviations are presented for continuous criterion variables; percentages are presented for categorical ones.

* $p < .05$,

** $p < .01$,

*** $p < .001$