

Psychological Methods

Taxometric Analysis as a General Strategy for Distinguishing Categorical From Dimensional Latent Structure

Robert E. McGrath and Glenn D. Walters

Online First Publication, January 16, 2012. doi: 10.1037/a0026973

CITATION

McGrath, R. E., & Walters, G. D. (2012, January 16). Taxometric Analysis as a General Strategy for Distinguishing Categorical From Dimensional Latent Structure. *Psychological Methods*. Advance online publication. doi: 10.1037/a0026973

Taxometric Analysis as a General Strategy for Distinguishing Categorical From Dimensional Latent Structure

Robert E. McGrath
Fairleigh Dickinson University

Glenn D. Walters
Federal Correctional Institution–Schuylkill

Statistical analyses investigating latent structure can be divided into those that estimate structural model parameters and those that detect the structural model type. The most basic distinction among structure types is between categorical (discrete) and dimensional (continuous) models. It is a common, and potentially misleading, practice to apply some method for estimating a latent structural model such as factor analysis without first verifying that the latent structure type assumed by that method applies to the data. The taxometric method was developed specifically to distinguish between dimensional and 2-class models. This study evaluated the taxometric method as a means of identifying categorical structures in general. We assessed the ability of the taxometric method to distinguish between dimensional (1-class) and categorical (2–5 classes) latent structures and to estimate the number of classes in categorical datasets. Based on 50,000 Monte Carlo datasets (10,000 per structure type), and using the comparison curve fit index averaged across 3 taxometric procedures (Mean Above Minus Below A Cut, Maximum Covariance, and Latent Mode Factor Analysis) as the criterion for latent structure, the taxometric method was found superior to finite mixture modeling for distinguishing between dimensional and categorical models. A multistep iterative process of applying taxometric procedures to the data often failed to identify the number of classes in the categorical datasets accurately, however. It is concluded that the taxometric method may be an effective approach to distinguishing between dimensional and categorical structure but that other latent modeling procedures may be more effective for specifying the model.

Keywords: categorical latent structure, dimensional latent structure, taxometric analysis, latent structural modeling, finite mixture modeling

An important distinction exists among statistical techniques available for the study of latent structure between those techniques used primarily to distinguish between dimensional versus categorical latent structure (the latent structure *type*) and those used to estimate a structural model assuming either a primarily dimensional or categorical structure (the latent structure *model*). Examples of techniques often used for the former purpose include finite mixture modeling and various statistics developed in conjunction with cluster analysis such as the cubic clustering criterion. Technically, these strategies identify dimensional structure by suggesting a one-class structure. Examples of techniques used for the latter purpose include traditional factor analytic and item response theory techniques, which assume the latent structure is purely dimensional, and cluster analysis, which assumes a categorical structure.

We would propose that some attempts to model latent structure without first evaluating whether the procedures used are consistent with the latent structure type are potentially misleading. For example, Moffitt's (1993) taxonomy of life-course-persistent and adolescent-limited offending has had a significant impact on research in delinquency and conduct disorder. However, preliminary research showing that the two might represent endpoints on a dimension rather than distinct types (Walters, 2011) suggested that scientific resources may be better spent investigating relative stability in delinquency dimensions than engaging in the more common practice of identifying distinct delinquency trajectories (Piquero, 2008). Furthermore, because latent structure holds important implications for research, theory, and practice (Ruscio, Haslam, & Ruscio, 2006), misidentification of latent structure type could lead to faulty selection of measurement models (groups vs. dimensions), faulty inferences about etiological process (specific etiology/nonlinear interaction/threshold effect/developmental bifurcation vs. additive models), and faulty development of assessment devices (using a small number of items designed to distinguish between categories vs. a larger number of items located at different points along a continuum).

This is not to say that every analysis of latent structure need begin with an evaluation of latent structure type. Before assessing latent structure type, the presence of categorical structure should be a reasonable hypothesis to consider. In a discussion of taxometric analysis, Lenzenweger (2004) listed three criteria that should be satisfied before undertaking an investigation into latent structure type: (a) a substantive and theoretically sound model is

Robert E. McGrath, School of Psychology, Fairleigh Dickinson University; Glenn D. Walters, Psychology Services, Federal Correctional Institution–Schuylkill.

The assertions and opinions contained herein are the private views of the authors and should not be construed as official or as reflecting the views of the Federal Bureau of Prisons or the United States Department of Justice. We are grateful to John Ruscio for his assistance with questions about the R program used in this study.

Correspondence concerning this article should be addressed to Robert E. McGrath, School of Psychology, Fairleigh Dickinson University, Teaneck, NJ 07555. E-mail: mcgrath@fdcu.edu

under investigation; (b) the model implicitly or explicitly posits the existence of a category or taxon; and (c) identifying a taxon will make a difference in how we understand or treat the construct. Walters (2011) applied these criteria to Moffitt's (1993) developmental taxonomy of antisocial behavior and discovered that the model was (a) coherent, substantive, and sound; (b) plausibly categorical; and (c) in a position to answer important questions about etiology, assessment, and treatment if, in fact, the latent structure of antisocial behavior were to prove categorical. Consequently, performing an analysis to determine latent structure type was justified.

The most basic distinction among latent structure types is the one between what we will refer to as *dimensional* and *categorical* structure types (Haertel, 1990).¹ The former can be further subdivided according to the number of dimensions, the latter according to the number of classes (recognizing that dimensional variation, and covariation, can also occur within classes). The most important subdivision within the categorical structure types is the distinction between dichotomous and polytomous (more than two classes) structure.

In recent years, taxometrics have emerged as a popular method for distinguishing between dimensional and dichotomous structure types.² Taxometrics were introduced by Meehl and his associates (e.g., Meehl, 1995, 2004; Meehl & Yonce, 1994, 1996; Waller & Meehl, 1998), who used the terms *taxon* and *complement* to refer to the two classes that make up a dichotomous structure. The emergence of taxometrics popularized the study of latent structure type, particularly in the fields of personality and psychopathology, and provided a statistical tool for such research.

Taxometrics have an important advantage over other structural typing strategies in that Meehl and his associates developed a variety of relatively independent taxometric procedures, each of which generally allows for multiple independent tests of the type of structure. Meehl (1995) coined the term *coherent cut kinetics* to refer to the use of consistency in findings across multiple distinct analyses as a guide to conclusions about the structure type. Three relatively independent taxometric procedures are in common use. These have been described in detail elsewhere and will only be summarized briefly here.³

Mean Above Minus Below a Cut (MAMBAC; Meehl & Yonce, 1994) requires two indicators of the construct. Cut scores are set at successive points on one variable (the input indicator). At each cut score, the mean score for a second indicator (the output indicator) is computed separately for those above and below the cut, and the difference between the two means is computed. For dichotomous constructs without serious skew, a graph with input cut scores on the abscissa and output mean differences on the ordinate should be hill shaped, with the peak occurring at the cut where the mean difference between the two classes is at a maximum. When the latent construct is dimensional, the same graph should tend toward a U shape. For all possible combinations of m variables used as both input and output indicators, MAMBAC allows $m(m - 1)$ distinct tests of structure type. It is also possible to generate an estimate of the taxon base rate (assuming a taxon exists) from each curve.

Maximum Covariance (MAXCOV; Meehl & Yonce, 1996) requires one input and two output indicators. Observations are ordered and divided into subsamples along the input indicator, and the covariance between the output indicators is computed within

each subsample. For dichotomous constructs, a graph with input subsample on the abscissa and output covariances on the ordinate should resemble a hill that peaks at the subsample where the frequencies of the taxon and complement classes are approximately equal. If the construct is dimensional, the graph should show no clear peak. With m indicators MAXCOV allows

$$\frac{m(m - 1)(m - 2)}{2}$$

distinct tests of structure type. Again, a separate estimate of the taxon base rate can be derived from each curve.

Latent Mode factor analysis (L-Mode; Waller & Meehl, 1998) requires a minimum of three indicators. In L-Mode the full set of indicators is subjected to factor analysis and a density plot of the factor scores for the first factor is generated. Because of the reduction in measurement error, the density plot for the first factor scores should be bimodal if the data reflect a dichotomous structure and unimodal if the data are dimensional. One L-Mode analysis is usually conducted using the entire set of indicators. Two estimates of the taxon base rate can be generated from the single L-Mode curve.

Although the number of distinct tests of structure type that can be generated from taxometrics offers a clear advantage over other tests of structure type, taxometrics have a significant disadvantage in that they were intended only to distinguish between dimensional and dichotomous structures. In contrast, cluster-analytic methods and finite mixture modeling can estimate the number of true classes whether the data are inherently dimensional, dichotomous, or polytomous.

Several publications have now considered the implications of polytomous structure for taxometric analysis. The first discussions of this issue were primarily conceptual, suggesting that taxometric analysis can be applied to such data so long as the categories are ordinally arranged on the manifest variables. Under these conditions, it was hypothesized that taxometric procedures could be used iteratively to identify one class boundary, then within the classes identified on the basis of the first set of analyses to identify subsequent boundaries (e.g., Meehl, 1999; Ruscio & Ruscio, 2002).

McGrath (2008) considered what would happen if taxometric procedures are applied to data with the intention of distinguishing between dichotomous and dimensional structure when the latent structure actually involves three classes. He found that standard statistics used to distinguish between dimensional and dichotomous structure produced confusing or even misleading results.

¹ Some authors refer to the former structure as *continuous*. We prefer the more general term *dimensional* because evidence that a latent structure is not categorical is insufficient evidence that the underlying distribution is continuous (Michell, 2005).

² Some articles have associated taxometric procedures with the *identification* of categorical structure, but for the present study it is particularly important to recognize that the taxometric method was developed specifically to *distinguish* between dimensional and dichotomous structure.

³ A fourth taxometric procedure, Maximum Eigenvalue (MAXEIG; Waller & Meehl, 1998), also frequently appears in the taxometric literature, but it is highly redundant with MAXCOV (Ruscio, Walters, Marcus, & Kacetow, 2010) and so will not be considered further here.

McGrath's (2008) findings highlighted the need to consider the possibility of polytomous structure when using taxometrics to evaluate questions of latent structure type. However, the iterative process described by Meehl (1999) and Ruscio and Ruscio (2002) has not yet been tested in practice. Complicating efforts to develop a generalization of taxometrics to the polytomous case was the absence of a standard criterion for drawing a conclusion about dimensional versus two-class structure from each of the taxometric tests.

Recently, the comparison curve fit index (CCFI; Ruscio, Ruscio, & Meron, 2007) has been suggested as such a standard. The CCFI is computed using the following steps:

1. For each of the taxometric procedures, compute the usual curves. Average the curves for taxometric procedures that produce more than one.
2. Create two comparison populations bootstrapped from the original data, with one population generated under the assumption of dimensional structure and one under the assumption of dichotomous structure. These distributions are intended to match the original data on distribution shape, correlations between indicators, and between-group differences in the taxonic case. Ruscio and Kaczetow (2008) developed a program that can be used to generate these comparison populations.
3. Draw multiple random samples from the two populations and submit these samples to the same taxometric analyses. Separately average the set of curves generated from the dichotomous and from the dimensional comparison populations.
4. Compute the root-mean-squared residual (RMSR) between corresponding data points on the mean curve generated from the actual data and the mean curve generated from the simulated dichotomous population. Do the same with the mean curve generated from the dimensional population.
5. To compute the CCFI, divide the RMSR for the dimensional simulations by the sum of the two RMSRs. The resulting statistic can range between 0 and 1, with a value below .50 consistent with dimensional structure and a value above .50 consistent with dichotomous structure. Whereas a cut score of .50 allows classification of all CCFI values as indicative of one structure or the other, Ruscio, Walters, Marcus, and Kaczetow (2010) showed that the statistic achieved higher accuracy rates when CCFI values in the range of .45 to .55, or .40 to .60, were treated as ambiguous.

The CCFI has now been found to be an effective indicator of dimensional or dichotomous structure in a number of studies (e.g., Ruscio et al., 2007, 2010; Walters & Ruscio, 2009, 2010). Recently, Walters, McGrath, and Knight (2010) demonstrated more generally that the CCFI accurately distinguishes between dimensional and categorical structures whether the latter are dichotomous or trichotomous. Their findings provide a rationale for evaluating the CCFI as a tool for implementing the iterative process

described by Meehl (1999) and Ruscio and Ruscio (2002). The present study presents the results from such an evaluation. Through iterative taxometric analyses using the CCFI as the criterion for gauging categorical structure, the results provide information about the use of taxometrics as a general strategy for distinguishing between dimensional and categorical structure, and for estimating the number of classes constituting a categorical structure.

Method

In this study we examined latent structure in Monte Carlo datasets reflecting dimensional ($k = 1$) and categorical ($k = 2-5$) structures. For each latent structure, 2,500 simulated datasets were generated with $m = 3-6$ continuous indicators, for a total of 10,000 datasets per structure type. These datasets were created using the iterative algorithm described by Ruscio and Kaczetow (2008). Their algorithm can be used to produce datasets reflecting dimensional structure that vary in sample size, number of indicators, target correlation between indicators, and asymmetry and tail weight of the indicator distributions. A dataset reflecting two-class latent structure, with a preset value for d and the variance ratio between the two classes, can be created by generating two subsamples of data, linearly transforming one subsample, and then combining the two. Applying the same strategy to k subsamples can be used to generate k -class datasets consisting of $k - 1$ taxon classes and a complement class (Walters et al., 2010). Note that this strategy generates classes that appear at the same relative location on each of the manifest variables, an issue we will return to later in the Discussion when addressing the limitations of the study.

Parameters for dimensional datasets were randomly selected from uniform distributions for sample size ($N = 1,000-3,000$), between-indicator correlations ($r_{xy} = .10-.65$), asymmetry ($g = .00-.30$), and tail weight ($h = .00-.15$). The use of the g and h formulas to create asymmetrical and peaked distributions from normally distributed data was described by Hoaglin (1985). For normal distributions, $g = h = 0$. The magnitude of g controls the asymmetry of the resulting distribution, the magnitude of h its kurtosis. The permitted values for g and h restricted the skew of any subsample to the range 0-3 and the kurtosis to the range 0-28.

Parameters for categorical datasets were set by selecting randomly from uniform distributions for the sample size ($N = 1,000-3,000$), taxon base rates ($BR = .10 - 1/k$, e.g., the maximum BR for any taxon class in a four-class dataset was 1/4), indicator validity ($d = 1.25-2.00$ between adjoining classes), taxon: complement variance ratio ($VR = 1-4$, integers only), within-group correlations ($r_{WG} = .00-.30$), asymmetry ($g = .00-.30$), and tail weight ($h = .00-.15$). Within polytomous datasets, BR and d were allowed to vary independently for each taxon class. The complement base rate represented the residual after summing the base rates for the taxon classes. A single value was selected for each of the within-class parameters (VR , r_{WG} , g , and h) for a dataset. The actual skew and kurtosis for each dataset were affected by the values for BR , d , and VR as well as g and h . The mean skew across all 50,000 datasets was .48 (range = $[-0.29, 1.72]$), and the mean kurtosis was 2.25 (range = $[1.67, 9.00]$).

Taxometric tests were computed using Ruscio's (2011) taxometric software. Datasets were analyzed using MAMBAC,

MAXCOV, and L-Mode. Analyses were conducted using default settings for the software at the time of its download, with the exception that MAMBAC curves were generated using a LOWESS smoother and MAXCOV curves using running medians (an additional variation will be noted below). MAXCOV subsamples were based on overlapping windows of cases, as recommended by Walters and Ruscio (2010). For purposes of computing the CCFI, 10 bootstrapped samples were drawn for each of the two scenarios.

Results

The analysis of the taxometric method as a strategy for distinguishing between latent dimensional and categorical structure proceeded in three stages. The first (Step 1) involved classifying datasets with one to five categories (where the presence of one category represents dimensional structure) as either dimensional or dichotomous on the basis of CCFI values. In those datasets where the results suggested dichotomous structure, base-rate estimates of the putative taxon were then used as a basis for dividing the dataset into two subsamples. Each subsample was then submitted to taxometric analysis with the purpose of identifying further categorical distinctions (Step 2). Those subsamples where the taxometric method suggested dichotomous structure were further subdivided, with each subsample submitted to taxometric analysis (Step 3). Step 1 permitted identification of one to two categories within each dataset, Step 2 permitted one to four categories, and Step 3 permitted one to eight categories.

Taxometric Analysis Step 1

Identifying basic structure. The first step in the iterative process involved submitting all 50,000 datasets to taxometric analysis and computing the CCFI value for each dataset. CCFI values $\leq .50$ were taken as support for dimensional structure, and values $> .50$ were taken as support for categorical structure. Five different standards for the detection of structure type were examined. These included the CCFIs generated separately by L-Mode, MAMBAC, and MAXCOV. In addition, the three CCFIs were averaged for each dataset, and the mean CCFI was compared with the same criterion. Finally, a majority vote procedure was investigated in which two CCFI values $\leq .50$ was considered evidence of dimensional structure, whereas two CCFI values $> .50$ was considered evidence of categorical structure. These last two were also evaluated by Ruscio et al. (2010).⁴

For each of the five detection standards, Table 1 provides six diagnostic efficiency statistics: sensitivity, specificity, positive predictive power, negative predictive power, overall hit rate, and area under the curve. The last provides an estimate of classification accuracy without reliance on a fixed cut score for the CCFI. Results are provided for the comparison of dichotomous ($k = 2$) versus dimensional ($k = 1$) datasets, which is consistent with previous research on the topic of classification using taxometrics. Results are also provided for the comparison of categorical ($k > 1$) versus dimensional ($k = 1$) datasets.⁵

Some of the results from the current study can be compared with results obtained by Ruscio et al. (2010) in a large-scale Monte Carlo investigation of the CCFI's ability to differentiate between dimensional and dichotomous datasets. Sensitivity values for de-

Table 1
Step 1 Diagnostic Efficiency Statistics for Distinguishing Dimensional Versus Dichotomous and Dimensional Versus Categorical Datasets

Structure	Sens	Spec	PPP	NPP	HR	AUC
L-Mode						
Dichotomous	81.1	100.0	100.0	84.1	90.5	1.00
Categorical	95.0	100.0	100.0	83.2	96.0	1.00
MAMBAC						
Dichotomous	99.4	99.8	99.8	99.4	99.6	1.00
Categorical	99.9	99.8	100.0	99.4	99.8	1.00
MAXCOV						
Dichotomous	89.2	99.9	99.8	90.2	94.5	1.00
Categorical	60.5	99.9	99.9	38.7	68.4	.99
Mean						
Dichotomous	96.5	100.0	100.0	96.7	98.3	1.00
Categorical	99.1	100.0	100.0	96.5	99.3	1.00
Majority						
Dichotomous	94.6	100.0	100.0	94.9	97.3	
Categorical	98.5	100.0	100.0	94.3	98.8	

Note. Dichotomous or Categorical structure was treated as the positive condition, and Dimensional structure was treated as the negative condition for classification analyses. Values are stated as percentages. Sens = sensitivity; Spec = specificity; PPP = positive predictive power; NPP = negative predictive power; HR = hit rate; AUC = area under the curve; L-Mode = Latent Mode comparison curve fit index (CCFI); Dichotomous = classification of 10,000 dichotomous versus 10,000 dimensional datasets; Categorical = classification of 10,000 dimensional versus 40,000 two- to five-class (categorical) datasets; MAMBAC = Mean Above Minus Below A Cut CCFI; MAXCOV = maximum covariance CCFI; M = classification based on the mean of the three CCFIs; Majority = classification based on the majority of the three CCFIs.

tecting dichotomous structure in our Table 1 correspond with values listed in Ruscio et al.'s Table 2 (p. 11) as the accuracy rate for categorical datasets at the .50 threshold. Our specificity values can be compared with the accuracy rate for dimensional datasets at the .50 threshold in their Table 2, and our hit rate parallels their accuracy rate for all samples.

There were some differences noted. In particular, L-Mode was less sensitive (81.1% vs. 93.1%) though more specific (100% vs. 96.3%) in our simulations than in the Ruscio et al. (2010) study. The discrepancies from Ruscio et al. may reflect differences in the value ranges for various parameters (particularly the number of predictors, sample sizes, and taxon base rates) across the two studies, as well as the influence of multiple taxonomic groups on the outcomes. Across the three taxometric procedures, however, the mean difference in sensitivity and specificity statistics was only 0.4%, whereas the mean difference in overall hit rate was 0.5%.

⁴ Ruscio et al. (2010) also investigated a unanimity standard that allowed classification only if all three CCFIs supported the same conclusion. However, preliminary results were not supportive of this approach, so to simplify the presentation it was omitted.

⁵ The statistics provided in Table 1 assume a rectangular distribution of the five structure types. Because the actual relative frequency of various structure types is unknown, this assumption could result in a mis-estimate of the CCFI's true accuracy. However, equalizing the proportion of categorical and dimensional datasets actually enhanced diagnostic efficiency estimates slightly since specificity tended to be better than sensitivity for detecting categorical datasets.

For the two aggregative approaches (Mean and Majority) the mean difference was only 0.2%, suggesting that the present results largely replicate previous findings.

On the basis of the results presented in Table 1, the decision was made to use the mean CCFI value as the criterion for categorical structure in subsequent analyses. This statistic accurately detected 96.5% of 10,000 dichotomous datasets, 99.1% of the 40,000 categorical datasets (which included the dichotomous sets), and 100% of the dimensional datasets. The overall hit rates for the MAMBAC CCFI alone were actually higher than those for the mean CCFI, suggesting that this value provided the best single indicator of structure type, but several factors argued in favor of using the mean. First, the difference was small and potentially unreliable given that the mean CCFI was associated with the higher hit rate in Ruscio et al. (2010). Second, the mean CCFI was associated with perfect prediction in the case of dimensional datasets, as indicated by specificity and positive predictive power values of 100%.

Additional analyses were conducted to identify predictors of accuracy in the CCFI. For two structural models (one and four classes), prediction using the mean CCFI as an index of dimensional or categorical structure was perfect. For each of the three structural models where there was variability in the outcomes for the dichotomized mean CCFI (two, three, and five classes), logistic regressions were conducted attempting to predict CCFI outcome using the variables N , base rate for the highest taxon group, base rate for the complement group, mean d value, VR , r_{WG} , g , h , and number of indicators. Results demonstrated little consistency across the three structural models, providing little guidance for predicting when the CCFI will be effective.

Estimating the taxon base rate. The other issue to be addressed with the Step 1 results was the optimal classification of cases into two classes once a dataset was identified as categorical based on CCFI results. Several studies suggest that the base-rate technique is the most accurate method for case classification available in taxometrics (Ruscio, 2009; Ruscio, Haslam, & Ruscio, 2006, pp. 146–148). This procedure involves assigning those observations with the highest total score across the indicators to the taxon class so as to match the estimated taxon base rate. Applying the base-rate technique, therefore, first requires generating an estimate of the taxon base rate from the dataset, so the next step in the analysis involved identifying an optimal base-rate estimation procedure.

Identifying an optimal base-rate estimation procedure is relatively straightforward for simulations with an underlying dichotomous structure, in that the various base-rate estimates generated from the taxometric analyses can be compared with the actual taxon base rate (e.g., Meehl & Yonce, 1994, Table 1; Meehl & Yonce, 1996, Table 2). The matter is more complicated in the case of polytomous datasets, because taxometric procedures for estimating base rates were developed only to distinguish between two classes. McGrath (2008) discussed the implications of trichotomous structure for taxometric base-rate estimation procedures and concluded that various parameters of the sample distribution can influence the results. Several possibilities could be justified on the basis of that discussion:

1. Taxometric procedures would tend to split the sample at the first taxon class, so that the complement class would

emerge as one class and all the taxon classes would be combined into the second.

2. Taxometric procedures would tend to split the sample into two relatively equal-sized classes.
3. Taxometric procedures would tend to split the sample where the indicator validity (the difference between two adjoining class means stated in terms of d) was greatest.

The goal was to identify the base-rate estimation procedure incorporated into taxometric software that came closest to detecting true dividing lines in the dataset. Several variables were created for each dataset that represented true dividing lines likely to be detected by existing taxometric procedures. The first was a series of base-rate values encompassing incremental numbers of taxon classes. For example, in datasets with five classes variable BR-4 was the base-rate estimate that resulted from treating the four taxon classes as one class and the complement class as the other; BR-3 was created by combining the three taxon classes with the highest means as one class and combining the complement class and lowest taxon class as the other; BR-2 was based on the highest two taxon classes versus the lowest three classes; and BR-1 was based on the highest taxon class versus the other four classes. Finally, variable BR- d was created by splitting the dataset at the taxon with the highest d value.

These variables were then compared with the mean base-rate estimates from L-Mode, MAMBAC, and MAXCOV. They were also compared with the mean of these three means (Mean3). Since the L-Mode mean is based on only two estimates derived from a single curve, and McGrath (2008) raised particular concerns about the impact of polytomous structure on L-Mode estimates of base rate, the dataset base rate values were also compared with the mean of the base-rate estimates from MAMBAC and MAXCOV (Mean2). Two sets of statistics were computed within each of the five structure types as the basis for comparisons: the correlation coefficient and the RMSR. The results of these analyses can be found in Table 2.

A number of conclusions can be drawn from this table. First, as predicted, both correlations and distance metrics indicated that L-Mode generated less accurate estimates of base rate than the other two taxometric procedures. The mean correlation with the dataset base-rate variables was actually negative, and the mean RMSR was substantially larger than that associated with the other procedures. These problems were evident even when datasets were dichotomous. Accordingly, subsequent discussion focuses on MAMBAC, MAXCOV, and Mean2.

Second, the remaining taxometric procedures were generally better at predicting dividing lines near the middle of the distribution. The mean correlations were higher for the more medial splits of the datasets, and the mean RMSRs were lower than was true for the base-rate variables in general, for those based on the largest d value, and for the division between taxon classes and the complement class. Finally, the Mean2 estimates nosed out those based on MAMBAC as predictors of the base rate derived from medial splits of the data. While Mean2 and MAMBAC were associated with the same mean RMSR (0.08), Mean2 was associated with a slightly larger mean correlation (.85). These findings suggest that the best single taxometric indicator for detecting a true dividing

Table 2
Step 1 Estimation of Base Rates

Class/ <i>M</i>	Correlation					RMSR				
	L-Mode	MAMBAC	MAXCOV	Mean3	Mean2	L-Mode	MAMBAC	MAXCOV	Mean3	Mean2
2 classes										
BR-1 ^a	0.00	0.66	0.97	0.87	0.89	0.30	0.09	0.04	0.10	0.06
3 classes										
BR-2 ^a	-0.05	0.88	0.91	0.79	0.92	0.20	0.13	0.13	0.07	0.13
BR-1 ^a	-0.43	0.77	0.78	0.43	0.80	0.38	0.11	0.11	0.19	0.11
BR- <i>d</i>	-0.23	0.66	0.75	0.50	0.73	0.31	0.10	0.09	0.14	0.10
4 classes										
BR-3	0.22	0.86	0.70	0.76	0.79	0.14	0.19	0.19	0.12	0.19
BR-2 ^a	-0.11	0.90	0.88	0.61	0.91	0.25	0.03	0.04	0.08	0.03
BR-1	-0.25	0.66	0.53	0.28	0.60	0.41	0.17	0.18	0.25	0.18
BR- <i>d</i>	-0.11	0.53	0.58	0.36	0.57	0.29	0.14	0.13	0.16	0.13
5 classes										
BR-4	0.27	0.84	0.48	0.67	0.65	0.13	0.24	0.24	0.17	0.24
BR-3 ^a	0.16	0.89	0.69	0.69	0.81	0.18	0.09	0.09	0.04	0.09
BR-2 ^a	-0.16	0.82	0.67	0.41	0.77	0.31	0.07	0.09	0.15	0.08
BR-1	-0.13	0.55	0.30	0.19	0.42	0.45	0.22	0.23	0.29	0.22
BR- <i>d</i>	-0.07	0.44	0.45	0.27	0.47	0.30	0.16	0.15	0.18	0.16
<i>M</i>	-0.07	0.73	0.67	0.52	0.72	0.28	0.13	0.13	0.15	0.13
<i>M</i> ^a	-0.10	0.82	0.82	0.63	0.85	0.27	0.08	0.08	0.11	0.08
<i>M-d</i>	-0.10	0.57	0.69	0.50	0.67	0.30	0.12	0.10	0.14	0.11
<i>M-k</i>	0.11	0.81	0.76	0.77	0.81	0.19	0.16	0.15	0.12	0.15

Note. Results for each structure type are based on 10,000 datasets. L-Mode = Latent Mode comparison curve fit index (CCFI); MAMBAC = Mean Above Minus Below A Cut CCFI; MAXCOV = maximum covariance CCFI; BR-(1-4) = base rate for the combined one to four taxon classes with the highest mean scores; BR-*d* = taxon base rate where the split is associated with the largest *d* value; *M*^a = mean for the rows with superscript ^a; *M-d* = mean of BR-*d* outcomes; *M-k* = mean based on the split between the complement and taxon classes; Mean3 = mean of all three mean base rate estimates; Mean2 = mean of MAMBAC and MAXCOV mean base rate estimates; RMSR = root-mean-squared residual.

^a Represents more medial splits of the dataset.

line between classes in the dataset is the mean of the mean base-rate estimates from MAMBAC and MAXCOV.⁶

Analyses were conducted to evaluate factors that predicted accuracy in estimation of the medial split of the data by Mean2. The absolute difference between the medial split of the data (averaging the two medial splits in cases where *k* was an odd number) and Mean2 was regressed onto the same nine variables used to predict accuracy in CCFI outcomes. This analysis was conducted separately for *k* = 2-5 classes. The following variables significantly predicted greater accuracy in at least three of the four analyses: larger BR for the highest class, larger BR for the complement class, larger mean *d* value, larger *r*_{WG} values, and smaller variance ratios. The direction of the relationship for *r*_{WG} was unexpected, as was the failure of *N* to prove a valid predictor. It should be noted that the relationship for *r*_{WG} was not significant in the dichotomous case, suggesting that it holds only for polytomous datasets, and the relatively large minimum value for *N* (1,000) may have suppressed any impact of sample size on the outcomes. That is, this finding should not be generalized to smaller sample sizes.

Taxometric Analysis Steps 2-3

To summarize the results from Step 1, the recommended single estimate of whether structure was dimensional or categorical was provided by the mean CCFI across all three taxometric procedures. Also, the mean of the mean base-rate estimates from MAMBAC and MAXCOV offered the optimal approach to identifying a true dividing line between classes within the dataset.

In Step 2, these heuristic rules were used as follows. Datasets identified as dimensional in Step 1 received no further attention. Datasets identified as categorical in Step 1 were divided into two subsamples based on the Mean2 estimate of base rate. Each of these subsamples was then evaluated using the three taxometric procedures.⁷

If neither subsample mean CCFI exceeded .50, the results were considered indicative of dichotomous structure (consisting of only two classes) and the procedure terminated at Step 2. If either or both subsample mean CCFIs exceeded .50, the process was repeated, dividing each subsample with mean CCFI > .50 into two subsamples based on Mean2, and the taxometric method was applied within each subsample (Step 3).

⁶ To evaluate whether Mean2 could be improved further as an estimate of the taxon base rate, the latter was regressed onto the former in the dichotomous samples and RMSR was computed between the predicted and actual base-rate values. This RMSR was only .002 smaller than the RMSR comparing the Mean2 and actual base-rate values.

⁷ The mean CCFI perfectly identified dimensional structure in this study. Since it is unlikely that all studies will achieve this level of accuracy, for purposes of "what if" analysis those dimensional datasets most likely to be identified as categorical—the 16 dimensional datasets that MAMBAC identified as categorical—were also submitted to Steps 2 and 3 analyses. Of those datasets, 15 were identified as consisting of only two classes, a fairly close estimate of the true structure.

To complete Step 3 analyses, two modifications of taxometric procedures were implemented. The Step 2 analyses identified as many as four subsamples of a dataset to be evaluated in Step 3. Even with $N > 1,000$ some subsamples were quite small. Subsamples with fewer than 100 cases were not analyzed further. Second, computation of the CCFI requires dividing the cases under analysis into two groups under the assumption of dichotomous structure. A sample-based estimate of the taxon base rate is used to determine the size of each group. In some subsamples one group was so small that the MAMBAC code crashed. Note that this is not a problem of subsample size per se but of the algorithm used to compute a statistic for that subsample. This problem was resolved by changing the MAMBAC code for the Step 3 analyses so that if the division of the subsample into two groups resulted in an error, MAMBAC was conducted again, this time dividing the subsample into two equal-sized groups.

Because each step allowed for doubling of the number of groups, it was possible after Step 3 for taxometric analysis to suggest anywhere from one to eight classes in the dataset. The distribution of outcomes across all datasets may be found in Table

3. Note that when all datasets were considered (so that small subsamples were treated simply as a single class), the mean estimate of the number of classes never differed from the true number by more than .45. The mean error declined further when only datasets that were completely analyzed (no small subsamples) were considered. As an overall index of consistency between true and estimated structure, the intraclass correlation (two-way mixed effects, single measure) was .87 for all datasets, .88 for those datasets where the analysis was not terminated at Step 3. However, while diagnostic efficiency statistics were generally high, some were disappointing. The positive predictive power when taxometrics suggested more than two classes, and the sensitivity for detecting three- or five-class structures, tended to be particularly poor. The latter finding probably reflects lower accuracy resulting from the averaged medial base-rate estimates when the number of classes was odd. When all datasets were considered, the overall hit rate for categorical datasets was only .44, though if mis-estimates by one class were allowed (e.g., if a three-class dataset was estimated to consist of either two, three, or four classes it was considered a hit), then the hit rate increased to .72. The values were

Table 3
Final Estimates of Number of Classes

Structure	Solution								<i>M</i>	Sens	Spec	Interp %
	1 class	2 classes	3 classes	4 classes	5 classes	6 classes	7 classes	8 classes				
All datasets												
1 class	10,000	0	0	0	0	0	0	0	1.00	1.00	0.99	100.00
2 classes	346	6,231	1,977	1,421	25	0	0	0	2.45	0.62	0.98	100.00
3 classes	2	855	4,939	3,633	476	95	0	0	3.40	0.49	0.94	100.00
4 classes	0	3	397	7,093	1,715	590	165	37	4.31	0.71	0.79	100.00
5 classes	12	0	84	3,390	3,862	1,483	816	353	5.06	0.39	0.94	100.00
PPP	0.97	0.88	0.67	0.46	0.64							
NPP	1.00	0.91	0.88	0.92	0.86							
Completely analyzed												
1 class	10,000	0	0	0	0	0	0	0	1.00	1.00	0.99	100.00
2 classes	346	6,231	1,269	854	14	0	0	0	2.31	0.72	0.98	87.14
3 classes	2	855	4,370	3,344	450	91	0	0	3.40	0.48	0.95	91.12
4 classes	0	3	394	7,021	1,696	589	165	37	4.31	0.71	0.80	99.05
5 classes	12	0	84	3,387	3,857	1,482	816	353	5.06	0.39	0.94	99.91
PPP	0.97	0.88	0.71	0.48	0.64							
NPP	1.00	0.94	0.89	0.91	0.85							
Unambiguous												
1 class	9,999	0	0	0	0	0	0	0	1.00	1.00	0.99	99.99
2 classes	134	4,755	456	369	0	0	0	0	2.19	0.83	0.99	57.14
3 classes	0	295	2,296	1,499	37	10	0	0	3.32	0.55	0.98	41.37
4 classes	0	0	101	4,719	328	59	10	2	4.07	0.90	0.86	52.19
5 classes	0	0	13	1,392	1,148	152	110	68	4.71	0.40	0.99	28.83
PPP	0.99	0.94	0.80	0.59	0.76							
NPP	1.00	0.96	0.93	0.97	0.93							
Unambiguous/completely analyzed												
1 class	9,999	0	0	0	0	0	0	0	1.00	1.00	0.99	99.99
2 classes	134	4,755	249	258	0	0	0	0	2.12	0.88	0.99	53.96
3 classes	0	295	1,995	1,431	35	10	0	0	3.33	0.53	0.98	37.66
4 classes	0	0	100	4,672	324	59	10	2	4.07	0.90	0.86	51.67
5 classes	12	0	13	1,391	1,144	152	110	68	4.69	0.40	0.99	28.90
PPP	0.99	0.94	0.85	0.60	0.76							
NPP	1.00	0.97	0.93	0.97	0.93							

Note. *Completely analyzed* refers to 47,722 (95.4%) datasets for which Step 3 analysis was not terminated due to insufficient sample size. *Unambiguous* refers to 27,952 (55.9%) datasets for which no tests proved ambiguous using ambiguity thresholds of [.45, .55]. *Unambiguous/completely analyzed* refers to 27,432 (54.9%) datasets meeting both conditions. Sens = sensitivity; Spec = specificity; Interp % = % of cases interpretable (not considered ambiguous or incompletely analyzed); PPP = positive predictive power; NPP = negative predictive power.

approximately the same when datasets with small subsamples were omitted. These findings suggest that although aggregating estimates of the number of classes across a large collection of datasets may achieve acceptable accuracy, the same is not true for estimates based on individual datasets.

As indicated previously, Ruscio et al. (2010) suggested several thresholds for interpretation of the CCFI, including treating CCFI values in the range [.45, .55] or [.40, .60] as ambiguous. Table 3 provides additional results using the former criterion. The inclusion of an ambiguous category created a cascading set of indeterminate outcomes across the three steps. In Step 1, the mean CCFI allowed for identification of structure in almost 97% of datasets, similar to the 99.4% identified by Ruscio et al. However, treating a dataset as ambiguous if the mean CCFI value fell in the more liberal interval [.45, .55] in any of the three steps dramatically reduced the proportion of interpretable outcomes, particularly for datasets with more categories. For the three- and five-class datasets the percentage of interpretable datasets fell below 50%, a finding consistent with the relatively poorer accuracy for those structures. Furthermore, though there was improvement in diagnostic efficiency statistics and the mean estimate of the number of classes, as could be expected, the increment would seem insufficient to compensate for the dramatic loss in information. Accordingly, we urge caution in the use of ambiguous intervals for CCFI when the goal is to estimate the number of categories in a polytomous structure.

Finite Mixture Modeling Comparison

To provide some context for the accuracy of the taxometric procedures, all datasets were also evaluated using finite mixture models as an alternative to taxometrics. Each dataset was evaluated for the presence of two to five classes. Analyses were conducted using Mplus (Muthén & Muthén, 2007). A recent comparison of some of the many fit indices Mplus can generate for mixture models suggested that the Bayesian Information Criterion (BIC) was the best of the information criteria examined, whereas the bootstrapped likelihood ratio test (BLRT) was the most consistent of the significance tests (Nylund, Asparouhov, & Muthén, 2007). Table 4 provides the mean BIC value across the 10,000 datasets reflecting each structural model as well as the probability that the p value for the BLRT was significant ($<.05$).

The BIC is relevant only to comparing categorical models. Smaller BIC values tend to be associated with better fit to the data. Similar to the results for the taxometric analyses, the BIC was biased toward models with more classes. Even if a decrease in $BIC < 100$ is treated as trivial, the results suggest a three-class solution for the datasets with one to three classes and a five-class structure for the datasets with four to five classes.

The BLRT can address questions of dimensional versus dichotomous structure as well as questions of the number of classes. Specifically, a significant outcome for the BLRT indicates that the corresponding model fits the data better than a model with one fewer class. A significant BLRT for the two-class model suggests the two-class model is superior to a one-class (dimensional) model, whereas a nonsignificant outcome suggests the two-class model is no better than the dimensional model. The proportion of datasets out of 10,000 in which the BLRT was significant is provided; a value $> .50$ for a model with k classes would suggest the k -class

Table 4
Results From Finite Mixture Modeling

Structure	<i>M</i> BIC	Δ BIC	$p(p < .05)$	Sens	Spec	PPP	NPP
1 class				.00	1.00	.00	1.00
2-class test	24,384.4		1.00				
3-class test	24,046.4	338.06	.75				
4-class test	23,976.6	69.73	.00				
5-class test	23,971.7	4.99	.00				
2 classes				1.00	.88	.67	1.00
2-class test	29,961.0		1.00				
3-class test	29,794.5	166.51	.00				
4-class test	29,749.1	45.41	.00				
5-class test	29,714.1	34.99	.00				
3 classes				.75	.89	.27	.89
2-class test	33,002.3		1.00				
3-class test	32,154.8	847.5	.75				
4-class test	32,062.9	91.92	.00				
5-class test	32,020.2	42.7	.00				
4 classes				.75	1.00	1.00	.94
2-class test	36,397.5		1.00				
3-class test	34,604.9	1792.56	1.00				
4-class test	34,233.3	371.59	.75				
5-class test	34,103.3	130.02	.00				
5 classes				.00	1.00	.00	.80
2-class test	39,114.6		1.00				
3-class test	36,408.3	2706.38	1.00				
4-class test	35,681.5	726.75	.00				
5-class test	35,296.1	385.38	.00				

Note. Datasets reflecting one to five classes were tested for two to five classes. Diagnostic efficiency statistics are based on bootstrapped likelihood ratio test results. *M* BIC = mean Bayesian Information Criterion; Δ BIC = change from previous model; $p(p < .05)$ = the probability of a p value $< .05$ across 10,000 datasets for the bootstrapped likelihood ratio test; Sens = sensitivity; Spec = specificity; PPP = positive predictive power; NPP = negative predictive power.

model fit the data better than the model with $k - 1$ classes. In each case, tests of the two-class model were always significant ($p = 1.00$), indicating consistent evidence of at least two classes rather than dimensional structure. The results suggested a three-class solution for both the one-class and five-class datasets. The diagnostic efficiency statistics provided are based on the BLRT's overall accuracy at detecting the number of classes.

Discussion

Research into latent structure tends to focus largely on structural modeling. Thousands of studies have been published in which some modeling technique, such as confirmatory factor analysis, item response theory, or cluster analysis, is applied to a set of data. Investigations into structure type that precede model estimation may in some cases be just as important (see also Markon & Krueger, 2006). Provided a case can be made for a construct's being potentially categorical (see Lenzenweger, 2004), the study of latent structure is optimally a two-stage process, in which the goal of the first stage is to determine the latent structure type (dimensional vs. categorical). Once the structure type is identified, the researcher can reasonably choose among the structural modeling techniques appropriate to that structure type to estimate the number of dimensions or classes and corresponding parameters (one might also recommend demonstration of both dimensional

structure and unidimensionality before the use of standard item response theory statistics).

The present study in combination with Walters et al. (2010) suggests that taxometrics can serve as a general strategy for distinguishing between dimensional and categorical structure types. In fact, the CCFI proved to be a better predictor of the type of structure in this study than finite mixture modeling. These findings are consistent with research suggesting that the taxometric method is particularly effective at differentiating between dimensional and dichotomous structure (e.g., Cleland, Rothschild, & Haslam, 2000; Frazier, Algorta, Youngstrom, & Ruscio, 2011) and that the mean CCFI is an accurate index of dichotomous structure (Ruscio et al., 2010).

Three limitations of the study are worth noting. The categorical structure represented here involved separation of classes along a single dimension. This is the condition for which the taxometric method is optimally suited as a technique for distinguishing between dimensional and categorical structure. However, categories can also appear as a set of locations on multiple dimensions. Other methods such as cluster analysis are more effective at detecting class structure under such circumstances. The researcher needs to evaluate the likely nature of the categorical structure before identifying an optimal procedure for attempting to distinguish categorical from dimensional structure. Similarly, dimensional datasets were consistently unidimensional in structure, which is the situation for which the taxometric method was designed. It is an empirical question yet to be addressed how well it would fare compared with alternatives in cases of multidimensional structure.

Second, the generalizability of our findings beyond the boundaries of the scenarios we studied is uncertain. Our datasets were limited to no more than five classes, because we suspect that data structures involving six or more groups are rare, but given that the mean CCFI achieved high levels of predictive power in the scenarios evaluated we would expect similar hit rates for even more complex categorical structures. The manifest variables we generated were multivalent, whereas many extant studies of latent structure use variables of limited range such as individual items as indicators. Walters and Ruscio (2009) found that indicators composed of four or more ordered levels were adequate for the purpose of discriminating between dimensional and dichotomous latent structures using MAMBAC and MAXCOV. Because the present study evaluated an iterative application of the taxometric method to that discrimination, even items of limited range are likely to produce adequate results. Another factor to consider in determining whether taxometric results are likely to be accurate is Meehl's (1995; see also Ruscio et al., 2010) recommendation to limit taxometric analyses to indicators with within-group correlations between indicators of $\leq .30$ and between-group separations of $d \geq 1.25$ (though these criteria are often impossible to ensure ahead of time in practice). Finally, the current study was restricted to samples of 1,000–3,000, which are relatively large. Ruscio et al. (2010) suggested samples of at least 500 when the latent structure is dimensional but samples of only 300 or less for dichotomous latent structure. We hypothesize, then, that samples as small as 500 may be adequate for detecting dimensional or categorical structure, but the minimum increases as the anticipated number of classes increases.

Third, the use of taxometrics to estimate the number of classes in a categorical latent structure can result in excessively high error

rates. However, this problem seems to be reduced if the outcome is averaged over a large series of datasets.

As noted earlier, some studies have evaluated treating values in the range of .45 to .55, or .40 to .60, as representing an ambiguous outcome for the CCFI. An ambiguous outcome by definition means the researcher cannot draw any conclusion about structure, and researchers in practice are unlikely to accept a decision-making algorithm that renders the results uninterpretable. In the current circumstances, we found that the application of an ambiguous interval in the context of a multistep analysis resulted in a large number of uninterpretable datasets. As an alternative, we would suggest that decisions based on values further from the natural cut score (.50) may merit a greater sense of subjective confidence in the decision than is true for those close to that value.

Once the taxometric method identifies a latent structure as categorical, other statistical methods may be superior at estimating parameters of the structural model such as the number and base rate of classes. Error rates for taxometric classification were small after Step 1 (see also Ruscio et al., 2010) but could be expected to increase with additional data analysis steps. On the other hand, the same should be true for any procedure that attempts to classify cases into four or six classes rather than two. There is very little literature currently available on the relative accuracy of different techniques for classifying cases into classes (see Beauchaine & Beauchaine, 2002).

Contrary to earlier speculations (e.g., Meehl, 1999; Ruscio & Ruscio, 2002), the iterative taxometric method is not an accurate approach to building a polytomous structural model at the level of the individual dataset. It is also somewhat complicated to implement (though R code for the conduct of taxometric analyses could easily be updated to perform iterative analyses when results suggest categorical structure), requires a large sample size, is greatly affected by base-rate inaccuracy, and can lead to conflicting results. A more serious issue is that it ultimately results in analyzing and dichotomizing uncomfortably small groups even when one starts out with a large sample size. We would conclude that the use of iterative taxometric procedures is optimal when used in conjunction with alternative procedures such as finite mixture models,⁸ recognizing that the taxometric method tends to overestimate the number of classes.

Despite the limitations noted in its use to determine the number of classes or class size in a categorical construct, taxometric analysis—using the mean CCFI as a criterion—may be a particularly accurate basis for decisions about dimensional versus categorical structure. We close with what we consider to be our most important point, which is that structural modeling should occur only after consideration of whether alternative hypotheses about the structure type are viable. Once structure type is identified, then the relevant structural modeling procedures can be applied with sound justification. The common practice of applying structural modeling techniques to data without a firm justification for assuming that the relevant structural type applies is unjustified and potentially misleading.

⁸ Cluster-analytic alternatives could also be mentioned as first-line indicators, but those methods can also be problematic (McGrath, 2008).

References

- Beauchaine, T. P., & Beauchaine, R. J. (2002). A comparison of maximum covariance and *k*-means cluster analysis in classifying cases into known taxon groups. *Psychological Methods*, *7*, 245–261. doi:10.1037/1082-989X.7.2.245
- Cleland, C. M., Rothschild, L., & Haslam, N. (2000). Detecting latent taxa: Monte Carlo comparison of taxometric, mixture model, and clustering procedures. *Psychological Reports*, *87*, 37–47.
- Frazier, T. W., Algorta, G. P., Youngstrom, E. A., & Ruscio, J. (2011). *Comparing taxometric analysis and latent variable models in the identification of categorical and dimensional data*. Manuscript submitted for publication.
- Haertel, E. H. (1990). Continuous and discrete latent structure models for item response data. *Psychometrika*, *55*, 477–494. doi:10.1007/BF02294762
- Hoaglin, D. C. (1985). Summarizing shape numerically: The *g*-and-*h* distributions. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Exploring data, tables, trends, and shapes* (pp. 461–513). New York, NY: Wiley.
- Lenzenweger, M. F. (2004). Consideration of the challenges, complications, and pitfalls of taxometric analysis. *Journal of Abnormal Psychology*, *113*, 10–23. doi:10.1037/0021-843X.113.1.10
- Markon, K. E., & Krueger, R. F. (2006). Information-theoretic latent distribution modeling: Distinguishing discrete and continuous latent variable models. *Psychological Methods*, *11*, 228–243. doi:10.1037/1082-989X.11.3.228
- McGrath, R. E. (2008). Inferential errors in taxometric analyses of ordered three-class constructs. *Journal of Personality Assessment*, *90*, 11–25. doi:10.1080/00223890701356755
- Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist*, *50*, 266–275. doi:10.1037/0003-066X.50.4.266
- Meehl, P. E. (1999). Clarifications about taxometric method. *Applied and Preventive Psychology*, *8*, 165–174. doi:10.1016/S0962-1849(05)80075-7
- Meehl, P. E. (2004). What's in a taxon? *Journal of Abnormal Psychology*, *113*, 39–43. doi:10.1037/0021-843X.113.1.39
- 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. doi:10.2466/pr0.1996.78.3c.1091
- Michell, P. (2005). *Measurement in psychology: A critical history of a methodological concept*. Cambridge, England: Cambridge University Press.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701. doi:10.1037/0033-295X.100.4.674
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*, 535–569. doi:10.1080/10705510701575396
- Piquero, A. R. (2008). Taking stock of developmental trajectories of criminal activity over the life course. In A. Liberman (Ed.), *The long view of crime: A synthesis of longitudinal research* (pp. 23–78). New York, NY: Springer-Verlag. doi:10.1007/978-0-387-71165-2_2
- Ruscio, J. (2009). Assigning cases to groups using taxometric results: An empirical comparison of classification techniques. *Assessment*, *16*, 55–70. doi:10.1177/1073191108320193
- Ruscio, J. (2011). *Taxometric programs for the R computing environment: User's manual*. Retrieved from <http://www.tcnj.edu/~ruscio/TaxProgManual%202011-11-13.pdf>
- Ruscio, J., Haslam, N., & Ruscio, A. M. (2006). *Introduction to the taxometric method: A practical guide*. Mahwah, NJ: Erlbaum.
- Ruscio, J., & Kacetow, W. (2008). Simulating multivariate nonnormal data using an iterative algorithm. *Multivariate Behavioral Research*, *43*, 355–381. doi:10.1080/00273170802285693
- Ruscio, J., & Ruscio, A. M. (2002). A structure-based approach to psychological assessment: Matching measurement models to latent structure. *Assessment*, *9*, 4–16.
- Ruscio, J., Ruscio, A. M., & Meron, M. (2007). Applying the bootstrap to taxometric analysis: Generating empirical sampling distributions to help interpret results. *Multivariate Behavioral Research*, *42*, 349–386. doi:10.1080/00273170701360795
- Ruscio, J., Walters, G. D., Marcus, D. K., & Kacetow, W. (2010). Comparing the relative fit of categorical and dimensional latent variable models using consistency tests. *Psychological Assessment*, *22*, 5–21. doi:10.1037/a0018259
- Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks, CA: Sage.
- Walters, G. D. (2011). The latent structure of life-course-persistent antisocial behavior: Is Moffitt's developmental taxonomy a true taxonomy? *Journal of Consulting and Clinical Psychology*, *79*, 96–105. doi:10.1037/a0021519
- Walters, G. D., McGrath, R. E., & Knight, R. A. (2010). Taxometrics, polytomous constructs, and the comparison curve fit index: A Monte Carlo analysis. *Psychological Assessment*, *22*, 149–156. doi:10.1037/a0017819
- Walters, G. D., & Ruscio, J. (2009). To sum or not to sum: Taxometric analysis with ordered categorical assessment items. *Psychological Assessment*, *21*, 99–111. doi:10.1037/a0015010
- Walters, G. D., & Ruscio, J. (2010). Where do we draw the line? Assigning cases to subsamples for MAMBAC, MAXCOV, and MAXEIG taxometric analyses. *Assessment*, *17*, 321–333. doi:10.1177/1073191109356539

Received July 24, 2010

Revision received October 18, 2011

Accepted November 11, 2011 ■