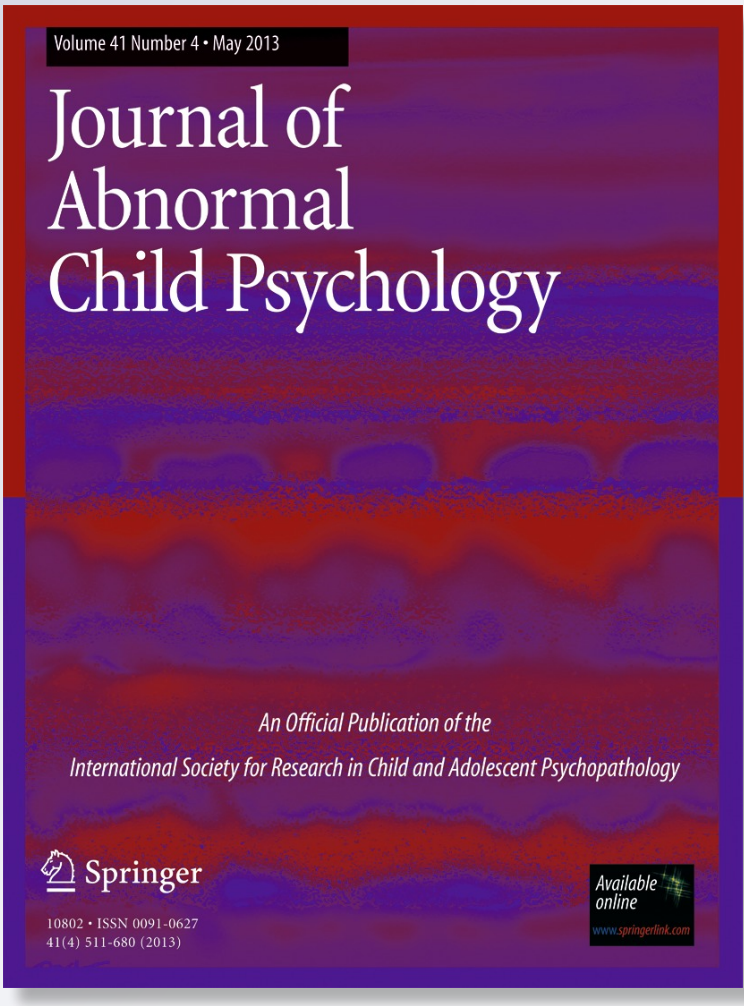
Trajectories of Youthful Antisocial Behavior: Categories or Continua?

Glenn D. Walters & John Ruscio

Journal of Abnormal Child Psychology
An official publication of the International Society for Research in Child and Adolescent Psychopathology

ISSN 0091-0627
Volume 41
Number 4

J Abnorm Child Psychol (2013) 41:653-666
DOI 10.1007/s10802-012-9700-1



Springer

J Abnorm Child Psychol (2013) 41:653–666
DOI 10.1007/s10802-012-9700-1

Trajectories of Youthful Antisocial Behavior: Categories or Continua?

Glenn D. Walters · John Ruscio

Published online: 28 November 2012

Abstract The purpose of this study was to determine whether qualitatively distinct trajectories of antisocial behavior could be identified in 1,708 children (843 boys, 865 girls) from the 1979 National Longitudinal Survey of Youth–Child Data (NLSY-C). Repeated ratings were made on the Behavior Problems Index (BPI: Peterson and Zill *Journal of Marriage and the Family*, *48*, 295–307, 1986) antisocial scale by the mothers of these children when the children were 6, 8, 10, 12, and 14 years of age. Scores on three indicators constructed from the six BPI Antisocial items (callousness, aggression, noncompliance) were then analyzed longitudinally (by summing across the rating periods) and cross-sectionally (by testing each individual rating period) in the full sample as well as in subsamples of boys and girls. Results obtained with the mean above minus below a cut (MAMBAC), maximum covariance (MAXCOV), and latent mode factor analysis (L-Mode) taxometric procedures revealed consistent evidence of continuous latent structure despite the fact Growth Mixture Modeling (GMM) and Latent Class Growth Analysis (LCGA) identified between two and eight trajectories, depending on the stopping rule, in the three antisocial indicators. From these results, it is concluded that the structural model underlying these data is better represented as continuous rather than as categorical. The implications of these results for future research on developmental trajectories of antisocial behavior are discussed.

The authors would like to thank anonymous reviewers of an earlier version of this paper for their helpful comments. Program code and a user's manual for the taxometric programs used in this study can be accessed at http://www.tcnj.edu/~ruscio/taxometrics.html.

G. D. Walters (✉)
Department of Criminal Justice, Kutztown University, Kutztown, PA 19530-0730, USA
e-mail: walters@kutztown.edu

J. Ruscio
Department of Psychology, The College of New Jersey, Ewing Township, NJ, USA

Keywords Taxometric method · Trajectories · Antisocial behavior · Latent structure

There is a growing belief among researchers and scholars that antisocial behavior in children conforms to a finite number of developmental trajectories. The precise number of developmental trajectories remains a matter of debate, but most studies show evidence of anywhere from two to seven trajectories or patterns of early antisocial or delinquent behavior. One of the more influential developmental models of antisocial behavior in children is the one advanced by Moffitt (1993). In her theory, Moffitt proposed a three-class taxonomy: life-course-persistent (LCP) antisocial behavior, adolescence-limited (AL) antisocial behavior, and a non-antisocial pattern. Whereas the LCP trajectory is marked by conduct problems beginning in early to middle childhood and persisting into middle adulthood, the AL trajectory is characterized by middle adolescent onset and desistance in late adolescence or early adulthood. Moffitt's (1993) theory has generated a substantial amount of research and controversy. As of July 28, 2012, her seminal 1993 paper had been cited 4,674 times in the professional literature (Publish or Perish, Harzing 2007). The principle controversy surrounding Moffitt's (1993) three types, however, has not been with their existence but with their number. Two to four trajectories were identified in a multi-site study by Broidy and colleagues (2003), and in other studies as many as five (Hoeve et al. 2008), six (Lacourse et al. 2002) and seven (Bushway et al. 2003) developmental trajectories have been identified.

The variability in the number of trajectories observed in research on antisocial development is cause for concern, but of even greater concern is the possibility that this variability stems from the methodology employed rather than from the latent structure of the antisocial construct. Semi-parametric group-based analysis (Nagin 1999) and growth mixture modeling (Muthén 2004) are two procedures commonly used to assess developmental trajectories of antisocial behavior. Although these procedures differ in several respects, both use a mixture of normal distributions to account for admixed non-normality. These procedures can produce misleading results, however, when non-normality is the result of skew or kurtosis in a continuous distribution rather than a consequence of class admixture (Bauer 2007; Bauer and Curran 2003). Semiparametric group-based and growth mixture modeling analysis introduce additional components into an admixed distribution to account for the non-normality. This, in turn, can lead to an overestimation of the number of classes in a distribution (McLachlan and Peel 2000).

Modeling techniques that do not automatically create components to account for non-normality may be helpful in testing whether a distribution of scores is continuous or categorical. One possibility is Meehl's (1995, 2004) taxometric method, which rests on a completely different set of assumptions than mixture modeling. Four core principles of taxometrics include: coherent cut kinetics, consistency testing, comparison curve analysis, and evidence-based procedures (Walters 2012). A cardinal feature of Meehl's approach is moving a cut score through a distribution of indicators (cut kinetics) to determine whether predictable results emerge (coherence). Four of the five taxometric procedures make use of coherent cut kinetic principles: mean above minus below a cut (MAMBAC: Meehl and Yonce 1994), maximum covariance (MAXCOV: Meehl and Yonce 1996), maximum eigenvalue (MAXEIG: Waller and Meehl 1998), and maximum slope (MAXSLOPE: Grove and Meehl 1993). The fifth procedure, latent mode factor analysis (L-Mode: Waller and Meehl 1998) is based on factor analytic principles rather than coherent cut kinetics.

Consistency testing is a second core principle of the taxometric method. Whereas many popular statistical procedures utilize the null hypothesis testing approach, Meehl (1995) emphasized consistency testing. Rather than rejecting the null hypothesis of no relationship or no effect if the probability is less than 5 % that the relationship or effect surfaced by chance, Meehl advocated for the convergence of non-redundant information, or consistency testing, in evaluating taxometric results. Even though MAMBAC, MAXCOV/MAXEIG, and L-Mode hold to a common formal-numeric definition of latent structure, each follows a somewhat different path in evaluating this latent structure (Ruscio and Ruscio 2004). When outcomes are consistent (continuous versus categorical) across non-redundant lines of evidence, we can have greater confidence in the results. Consistency testing also entails consistency between the individual MAMBAC and MAXCOV/MAXEIG curves, in that each procedure produces multiple curves, and consistency across subsamples of participants selected on the basis of gender, race, or some other variable.

Meehl (1992, 1995) originally designed the taxometric method to test whether certain constructs of interest to him (e.g., schizotypy) could be classified as taxonic or categorical. Because the goal was to detect taxa, results inconsistent with categorical latent structure were labeled nontaxonic. Later, Meehl (2004) clarified a three-outcome alternative to the taxon-detection framework: results can consistently support categorical structure, consistently support continuous structure, or be ambiguous. Along these lines, Ruscio et al. (2006) proposed an inferential framework in which categorical and continuous latent structure were conceptualized as competing structural hypotheses. Ruscio et al.'s (2007) comparison curve approach was a logical extension of this new framework. By creating comparison curves from simulated categorical and continuous data and fitting these comparison curves to the data curve it was possible to assess the relative performance of the two competing hypotheses (categorical and continuous). The comparison curve fit index (CCFI) is calculated with the root mean squared residual (RMSR) of the fit between the categorical comparison curve and the actual data curve and the RMSR of the fit between the continuous comparison curve and actual data curve: $\mathrm{Fit}_{\text{RMSR-dim}} / (\mathrm{Fit}_{\text{RMSR-dim}} + \mathrm{Fit}_{\text{RMSR-tax}})$. CCFIs above 0.50 are most consistent with categorical latent structure, CCFIs below 0.50 are most consistent with continuous latent structure, and CCFIs around 0.50 are considered ambiguous or indeterminate (Ruscio et al. 2007).

The fourth core principle of taxometrics is that the non-redundant sources of information used in consistency testing are evidence-based or empirically-verified. Goodness of fit (GFI) estimates calculated between the predicted and obtained variance–covariance matrices, consistency between base rates obtained from different procedures and from different combinations of indicators using the same procedure, and curve shape have traditionally been used to evaluate taxometric results. Unfortunately, none of these procedures has been supported by evidence from rigorous study (Cleland et al. 2000; Haslam and Cleland 2002; Ruscio 2007; Ruscio et al. 2006). CCFI values calculated from comparison categorical and continuous curves have yielded consistently positive results, however, in ten separate Monte Carlo studies (McGrath and Walters 2012; Ruscio 2007; Ruscio et al. 2007, 2010; Ruscio and Marcus 2007; Ruscio and Walters 2009, 2011; Walters et al. 2010; Walters and Ruscio 2009, 2010). A study of consistency testing suggests that CCFI results for MAMBAC, MAXCOV, and L-Mode be averaged and values around 0.50 be considered indeterminate (Ruscio et al. 2010); specifically, CCFI $\leq$ 0.45 supports continuous structure, CCFI $\geq$ 0.55 supports categorical structure, and 0.45 < CCFI < 0.55 is ambiguous. Even though the taxometric method compares only two kinds of structure (categorical, continuous), research indicates that it is capable of identifying categorical latent structure in distributions composed of anywhere from two to five latent categories (Frazier et al. 2011; McGrath and Walters 2012; Walters et al. 2010).

Walters (2011) used taxometrics to assess whether Moffitt's (1993) taxonomic theory of antisocial behavior was a true taxonomy. In this study, maternal ratings of childhood antisociality, hyperactivity, and headstrongness were analyzed in participants who subsequently reported delinquent behavior between the ages of 15 and 21. The results failed to show evidence of discrete categories of early behavioral problems (a major feature of Moffitt's theory) and raised questions about the taxonomic status of Moffitt's three-class model. Moreover, the continuous results suggested that individual differences in early behavioral problems were better modeled as continua than as discrete classes of behavior. Whereas the Walters (2011) investigation examined the early behavioral deviance component of Moffitt's (1993) theory, it left untouched the developmental or longitudinal aspect that lies at the heart of her three-class model. A core assumption of Moffitt's (1993) theory is that LCP and AL delinquents display widely divergent patterns of delinquent behavior: LCP delinquents are said to have an earlier age of onset and longer course while AL delinquents have a later age of onset and shorter course. This assumption could not be tested in the Walters (2011) investigation because the indicators were cross-sectional rather than longitudinal.

Few attempts have been made to apply the taxometric method to longitudinal or time series data. It may nevertheless be possible to analyze longitudinal data taxometrically by summing scores across the time series and performing a taxometric analysis of the summed scores. Difference scores between adjacent time periods were initially considered for this purpose but because difference scores compound error variance across measures (Cronbach and Furby 1970), summed scores were used instead. Provided the putative trajectories differ primarily in magnitude (more or less antisociality) rather than in direction (positive or negative slope), this approach should be conceptually similar to the difference score approach but with less error and greater precision. It is also possible that categories may exist at one age but not another and so it is important to conduct analyses at each age. The purpose of this study was to determine whether childhood antisocial behavior was categorical or continuous. Using data from a large longitudinal sample, the current study subjected summed scores from a time series and cross-sectional data from each time period to taxometric analysis. It was hypothesized that the taxometric method would support a conclusion of continuous latent structure despite the fact growth mixture modeling and latent class growth analysis showed evidence of two or more trajectories.

Method

Participants

The National Longitudinal Survey of Youth 1979-Child Data (NLSY-C: Center for Human Resource Research 2009) is comprised of children born to females from the nationally representative NLSY79 sample. As of 2006, a total of 11,466 children had been identified as offspring of the original 6,283 female NLSY79 respondents. A subsample of 1,708 children born between 1980 and 1994 had been rated by their mothers on the Behavior Problems Index (BPI: Peterson and Zill 1986) every other year between the ages of 6 and 14. These individuals served as participants in the current study. The gender breakdown for the current sample was 843 males (49.4 %) and 865 females (50.6 %). Over half the sample was white (56.9 %) and the remainder of the sample was composed of black (27.0 %) and Hispanic (16.1 %) respondents.

Indicators

Indicators for this study were based on biennial maternal ratings of the Behavior Problems Index (BPI: Peterson and Zill 1986). The BPI is a structured interview designed to assess various internalizing and externalizing problems of childhood. The 28 BPI items were rated on a three-point scale for the NLSY-C: *often true* (2), *sometimes true* (1), and *not true* (0). These 28 items are organized into six factor scores (antisocial, anxious/depressed, headstrong, hyperactive, dependent, peer conflict/withdrawal). Research suggests that the BPI generates scores with adequate internal consistency, test–retest reliability, and concurrent validity (Baker et al. 1993). The six items from the antisocial factor were used to create indicators for this study. Because taxometric analysis performs best with quasi-continuous indicators of four or more ordered categories (Walters and Ruscio 2009) and some taxometric procedures require at least three indicators, the six items were paired based on their correlations with each other. Out of a total of 15 correlations between the six summed items, Items 5 (*Is disobedient at school*) and 6 (*Has trouble getting along with teachers*) correlated the highest (r=0.73), followed by the correlation between Items 2 (*Bullies or is cruel/mean to others*) and 4 (*Breaks things deliberately*), r=0.54. The correlation between Items 1 (*Cheats or tell lies*) and 3 (*Does not feel sorry after misbehaving*), r=0.47, though not the third highest, was above the median (0.46). Based on these correlational patterns, Items 1 and 3 were combined to form a *Callousness* indicator, Items 2 and 4 were combined to form an *Aggression* indicator, and items 5 and 6 were combined to form a *Noncompliance* indicator.

Procedure

The principal difference between Nagin's (1999) semi-parametric group-based growth analysis approach and Muthén's (2004) growth mixture modeling (GMM) approach is that Muthén's method allows for within group variability, whereas Nagin's method assumes that within-group variability is random error. Muthén's (2004) latent class growth analysis (LCGA) procedure is similar to Nagin's semi-parametric group-based growth analysis procedure in that both are special cases of GMM in which the variance and covariance estimates for the growth factors are fixed at zero within each class. In the current study, GMM for continuous data (with gender serving as a covariate) and LCGA for continuous data were performed on the antisocial indicators (callousness, aggression, noncompliance) across the five rating periods (ages 6, 8, 10, 12, and 14) using algorithms from *MPlus 5.0* (Muthén and Muthén 1998–2007). Maximum likelihood with robust standard errors (MLR) was employed as the estimator and the growth curves were specified with higher order polynomials. There were 200 random starts during the initial stage of each analysis and 20 optimizations during the final stage of analysis. Starting with a one-class model, an additional class was added to the model in a stepwise fashion until the stopping rule was achieved. Two different stopping rules were used in this study. The first stopping rule was a nonsignificant ($p \geq 0.05$) Lo-Mendell-Rubin adjusted likelihood ratio test (LMR: Lo et al. 2001) and the second stopping rule was an increase in both the Akaike information criterion (AIC) and Bayesian information criterion (BIC).

MAMBAC, MAXCOV, and L-Mode were computed using Ruscio's (2011) taxometric programs for R language (R Development Core Team 2005). Ordering the input indicator along the x-axis, MAMBAC calculates the mean score of the output variable above and below successive equally-spaced cuts (50 in the current study) along the input variable and plots this on the y-axis. Categorical constructs typically display a peak on the MAMBAC curve whereas continuous constructs show evidence of a dish-shaped curve that turns up at one or both ends (Meehl and Yonce 1994). Tied scores on the input indicator were resorted ten times, analyses were re-run, and results were averaged to stabilize the curves. Cases were assigned to the putative taxon and complement groups with the base rate classification procedure (Ruscio 2009). In this study a two-variable composite served as the input indicator and the third variable served as the output indicator in all possible combinations, thus yielding three curves. These curves were subsequently averaged to calculate a fit index (see below).

With MAXCOV, cases are sorted on the input indicator and grouped into ordered subsamples along the x-axis. The covariance between two output indicators is then computed for each subsample and plotted along the y-axis. Categorical constructs characteristically yield a peak in the subsample or subsamples that divide categories because covariation is highest under conditions of maximum heterogeneity (i.e., roughly equal mix of taxon and complement members). The MAXCOV curve exhibits no discernible peak when the construct being evaluated is continuous (Meehl and Yonce 1996). In the current study, the input indicator was organized into 25 sliding windows with 90 % overlap. Tied scores on the input indicator were resorted ten times to minimize their obfuscating effect. Cases were assigned to the putative taxon and complement groups using the base rate classification procedure. Arranging the three indicators into all possible triplet combinations (one input indicator and two output indicators) produced three curves which were then averaged and analyzed.

The logic behind the L-Mode procedure is that by plotting the distribution of participants' scores on the first principal factor it should be possible to distinguish between results for continuous (unimodal) and categorical (bimodal) latent structure (Waller and Meehl 1998). The L-Mode procedure involves extracting the first (and largest) principal factor of the modified indicator covariance matrix (with variances along the diagonal replaced by 0s) and plotting the distribution of participants' scores on this single latent factor using Bartlett's (1937) method of factor score estimation. Because the taxon base rate estimates provided by L-Mode can be less accurate than those of other procedures (Ruscio and Walters 2009), the MAMBAC and MAXCOV base rate estimates were averaged and used to assign cases to putative groups for all three procedures (MAMBAC, MAXCOV, L-Mode).

To the extent that deviations from normality (e.g., skew) can greatly disrupt the classic categorical and continuous patterns previously described, the MAMBAC, MAXCOV, and L-Mode curves were compared to curves derived from simulated categorical and continuous comparison data (Ruscio et al. 2007). The relative fit of each comparison curve to the actual data curve was computed using the root mean square residual (RMSR) of fit between the data curve and the two comparison curves: $CCFI = FIT_{RMSR\text{-}Dim} / (Fit_{RMSR\text{-}Dim} + Fit_{RMSR\text{-}Cat})$. A CCFI of 0.00 indicates maximum support for the continuous model, a CCFI of 1.00 indicates maximum support for the categorical model, and a CCFI of 0.50 indicates equal support for the continuous and categorical models. The dual-threshold criterion described earlier was used in this study: $CCFI \leq 0.45$ supports continuous structure, $CCFI \geq 0.55$ supports categorical structure, $0.45 < CCFI < 0.55$ is ambiguous.

Based on the results of the GMM and LCGA analyses, artificial categorical data were constructed for comparison to the actual data obtained in this study. The artificial categorical data were created by assigning cases to two groups (trajectories). For each group, a finite population with 100,000 times that group's base rate, was generated. Ruscio's (2011) GenData program reproduces distributions using a bootstrapping technique (thereby duplicating distributional moments) that then reproduces correlations using an iterative algorithm that hones in on the empirical data's correlation matrix. Finally, the two groups' data are merged to obtain a single set representing a finite population with N=100,000 cases (Ruscio and Kaczetow 2008). From this population of data, 100 random samples of $N=1{,}708$ were drawn and submitted to taxometric analysis. This same procedure was used to create artificial categorical data divided into two groups at the boundary between low and high/moderate elevation in the LMR- and AIC/BIC-identified solutions for each of the three antisocial indicators (callousness, aggression, noncompliance).

Results

Latent Growth Analyses

GMM and LCGA were applied to the antisocial indicators (callousness, aggression, noncompliance) across the five rating periods (ages 6, 8, 10, 12, and 14). Using the LMR stopping rule, both GMM and LCGA identified a three-class solution for callousness, a two-class solution for aggression, and a two-class solution for noncompliance (see Table 1). Using the AIC/BIC stopping rule, GMM identified a seven-class solution for callousness (although only five classes contained cases), a three-class solution for aggression, and an eight-class solution for noncompliance, whereas LCGA identified a five-class solution for callousness, an eight-class solution for aggression, and a six-class solution for noncompliance (see Table 2). Posterior probability class assignments obtained with GMM and LCGA correlated 0.99 for callousness, 1.00 for aggression, and 0.98 for noncompliance when the LMR stopping rule was used, and 0.87 for callousness, 0.74 for aggression, and 0.86 for noncompliance when the AIC/BIC stopping rule was used. GMM-derived trajectories for the three antisocial indicators created with the LMR and AIC/BIC stopping rules are reproduced in Figs. 1 and 2, respectively. Trajectories of antisocial behavior differed widely in magnitude but, with the exception of the GMM six-class model for noncompliance, not in direction (crossing trajectories).

Pre-Taxometric Analyses

It is important that the data meet certain requirements if taxometrics is to prove useful. Meehl (1995) recommended a total sample size of at least 300 and a putative taxon base rate of at least 10 %. With a total N of 1,708 the current sample clearly meets the sample size criterion. The fact that GMM-identified low elevation antisocial trajectories (average score in each of the five rating periods <1.00) predominated over moderate to high elevation antisocial trajectories (67.6 % low elevation in LMR-callousness, 87.2 % low elevation in LMR-aggression, 88.6 % low elevation in LMR-noncompliance, 63.9 % low elevation in AIC/BIC-callousness, 73.1 % low elevation in AIC/BIC-aggression, and 74.1 % low elevation in AIC/BIC-noncompliance) makes for a putative taxon base rate of 24.2 % (100 %−[67.6+87.2+88.6+63.9+73.1+74.1/6]). Meehl also recommends that indicators distinguish between the putative taxon and complement with an effect size of at least $d=1.25$ and that indicators correlate no higher than $r=0.30$ within each of the two putative groups. The present data appear to meet these requirements as well.

To estimate the data parameters, cases were assigned to groups such that the highest-scoring 24.2 % of cases were assigned to the putative taxon and the remaining 75.8 % of cases were assigned to the complement. Data parameters were also estimated using the base rates identified by the actual taxometric procedures. Across all six indicator sets, no indicator failed to distinguish between putative groups with $d<1.25$, nor was the mean inter-indicator correlation ever larger than $r=0.30$ within a putative taxon or complement group, except for the mean complement inter-indicator correlation for the summed indicators when the taxometric-generated base rate was used (see Table 3). In addition, the mean correlation between indicators in the full sample exceeded Meehl's recommended threshold of $r\geq 0.30$ in all six indicator sets. Correlations for the three indicators across the five age periods can be found in Table 4.

Taxometric Analyses

Analyses began with the summed (longitudinal) indicator set and continued with the five cross-sectional data sets, one for each age group in which the BPI was administered. Each taxometric analysis was accompanied by the generation of populations of categorical and dimensional (continuous) comparison data, with 100 random samples drawn from each population for parallel analysis to calculate a CCFI value. Table 5 lists the taxon base rate estimates and CCFI results for the 6 data sets in the total sample ($N=1{,}708$), the 6 data sets in the male subsample ($n=843$), and the 6 data sets in the female subsample ($n=865$). Seventeen of the 18 mean CCFI values were below 0.45, supporting an inference of continuous latent structure. The one indeterminate result (8 year old-female subsample) fell just short of the 0.45 cutoff for continuous structure. Figure 3 depicts the MAMBAC, MAXCOV, and L-Mode data curves relative to their respective categorical and continuous comparison curves. The general consistency of results across the three taxometric procedures, six indicator sets, and in the full sample and gender-specific subsamples provides strong evidence in favor of continuous latent structure.

Analyses of Artificial Data

Two-group artificial or comparison categorical curves were created using a GMM-identified taxon base rate

Table 1 Summary of GMM results for three indicators measured over five rating periods

Classes	Callousness					Aggression					Noncompliance				
	k	LL	AIC	BIC	LMR	k	LL	AIC	BIC	LMR	k	LL	AIC	BIC	LMR
One	12	−10,348	20,720	20,786	–	12	−7,612	15,248	15,314	–	12	−8,321	16,667	16,732	–
Two	16	−10,182	20,396	20,483	0.000	16	−7,222	14,476	14,563	0.000	16	−7,862	15,756	15,843	0.042
Three	20	−10,133	20,267	20,377	0.025	20	−6,284	12,608	12,716	0.539	20	−6,995	14,030	14,139	0.626
Four	24	−8,861	17,770	17,901	0.110	24	−6,284	12,616	12,747	0.500	24	−5,140	10,329	10,460	0.559
Five	28	−8,852	17,760	17,912	0.000						28	−5,032	10,121	10,274	0.050
Six	32	−8,454	16,972	17,146	0.999						32	−4,932	9,928	10,102	0.152
Seven	36	−8,231	16,534	16,730	0.999						36	−4,921	9,914	10,110	0.419
Eight	40	−8,852	17,784	18,002	0.999						40	−4,755	9,590	9,808	0.295
Nine											44	−4,755	9,598	9,837	0.000

The number of random starts was set at 200 during the initial stage and the number of optimizations was set at 20 during the final stage; *GMM* growth mixture modeling; *Class* number of classes; *k* number of free parameters; *LL* log-likelihood; *AIC* Akaike information criterion; *BIC* Bayesian information criterion; *LMR* probability of the Lo-Mendall-Rubin adjusted LRT test

of 0.242. This was calculated by averaging the base rates of high/moderate elevation versus low elevation for the three indicators and two stopping rules (six base rates in all). Using a GMM-identified average taxon base rate of 0.242, CCFIs of 0.215, 0.148, and 0.337 were obtained for MAMBAC, MAXCOV, and L-Mode, respectively. When a LCGA-identified taxon base rate of 0.270 was employed, the CCFIs for MAMBAC, MAXCOV, and L-Mode were found to be 0.207, 0.146, and 0.340, respectively. In order to save space, only the two-group categorical comparison curves created using the GMM-identified taxon base rate are shown in Fig. 4, although the results achieved with the LCGA-identified taxon base rate were nearly identical to those obtained using the GMM-identified base rate. Both sets of results were more consistent with continuous latent structure than with a two-class categorical model of latent structure.

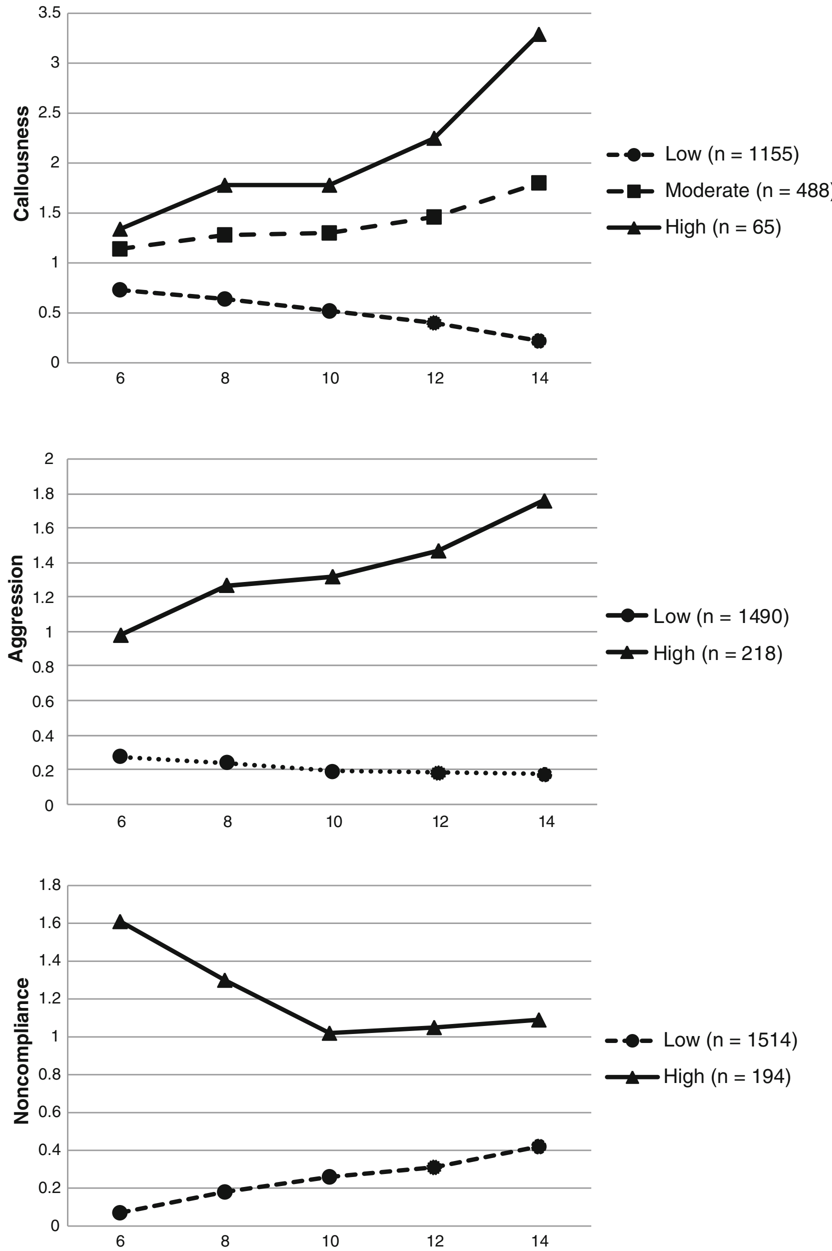
The artificial categorical data created when a boundary was inserted between the GMM- and LCGA-generated low and high/moderate elevation groups revealed a markedly different taxometric pattern than the pattern observed in the empirical data. The categorical MAMBAC comparison curves displayed a slight peak or incline near the right end of the curve that was not replicated in the actual data, and the categorical MAXCOV comparison curves exhibited a distinct peak to the right of center that was clearly not reflected in the actual data. There were two differences between the categorical L-Mode comparison curves and actual data curve: (1) the larger bump in the actual data curve was

Table 2 Summary of LCGA results for three indicators measured over five rating periods

Classes	Callousness					Aggression					Noncompliance				
	k	LL	AIC	BIC	LMR	k	LL	AIC	BIC	LMR	k	LL	AIC	BIC	LMR
One	10	−10,355	20,731	20,786	–	10	−7,630	15,280	15,334	–	10	−8,375	16,770	16,824	–
Two	13	−10,189	20,405	20,476	0.000	13	−7,240	14,506	14,577	0.000	13	−7,916	15,857	15,928	0.041
Three	16	−10,122	20,275	20,362	0.021	16	−7,038	14,109	14,196	0.212	16	−7,045	14,122	14,210	0.624
Four	19	−8,868	17,775	17,879	0.112	19	−6,302	12,642	12,745	0.499	19	−5,190	10,419	10,522	0.546
Five	22	−8,859	17,763	17,883	0.002	22	−6,112	12,269	12,389	0.276	22	−5,082	10,210	10,329	0.040
Six	25	−8,864	17.779	17,916	0.969	25	−4,404	8,859	8,995	0.036	25	−4,983	10,015	10,151	0.132
Seven						28	−4,203	8,461	8,614	0.072	28	−4,983	10.021	10,173	0.552
Eight						31	−4,180	8,422	8,591	0.188					
Nine						34	−4,405	8,877	9,062	0.113					

The number of random starts was set at 200 during the initial stage and the number of optimizations was set at 20 during the final stage; *LCGA* latent class growth analysis; *Class* number of classes; *k* number of free parameters; *LL* log-likelihood; *AIC* Akaike information criterion; *BIC* Bayesian information criterion; *LMR* probability of the Lo-Mendall-Rubin adjusted LRT test

Fig. 1 Mean indicator scores for participants in LMR-identified GMM trajectory groups



slightly shorter in magnitude than the larger bump in the categorical curves, and (2) the right side of the actual data curve was relatively smooth compared to a small but distinct second bump on the right side of the categorical L-Mode comparison curves. The fact that categorical data constructed from base rates identified by GMM and LCGA differed substantially from those for the empirical data suggests that the taxometric method was capable of uncovering evidence of discontinuity if, in fact, discontinuity existed. Because no evidence of discontinuity emerged, and because of the especially good fit between the results for empirical data and continuous comparison data, it would appear that the structural model underlying these data is best represented as continuous rather than as categorical.

Discussion

In an attempt to assess the longitudinal or time series features of Moffitt's (1993) and others' (Nagin et al. 1995; Nagin and Tremblay 1999) developmental theories of antisocial behavior, the sum of ratings over five rating periods and each of the five individual rating periods for three antisocial indicators were analyzed with the taxometric method. From the results of this study it would appear that the latent structure of antisocial behavior, as evaluated by the mothers of the children in the NLSY-C sample, is continuous rather than categorical. There was no evidence of discontinuity in the underlying distribution of data from the current study to support the existence of one or more

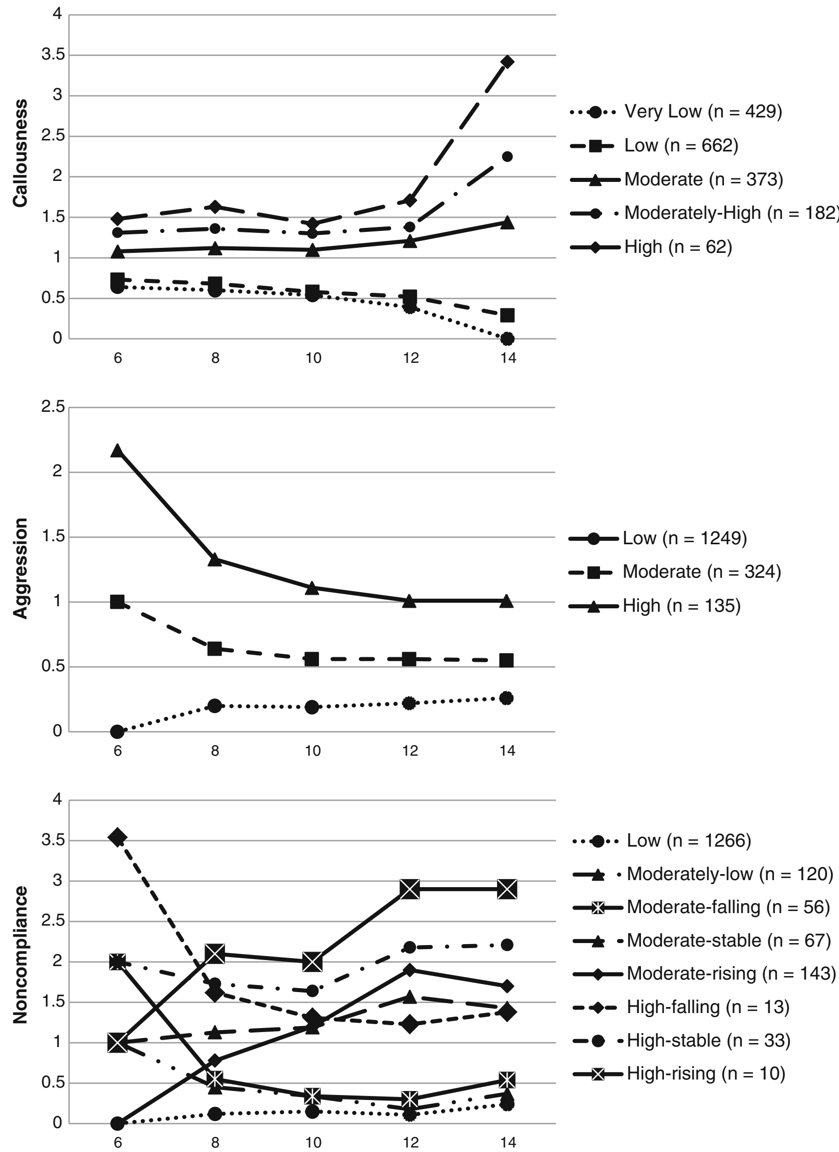


Fig. 2 Mean indicator scores for participants in the AIC/BIC-identified GMM trajectory groups

qualitatively distinct antisocial trajectories. This, along with the results of several other studies (Burt et al. 2011; Lahey et al. 2006; Walters 2011), lends credence to many of the criticisms that Sampson and Laub (2005) have leveled against the popular notion that a set of distinct developmental trajectories with divergent etiological processes and pathways exist for delinquency and antisocial behavior. Studies on GMM, in fact, indicate that one-category solutions are rarely identified by this procedure, even when the underlying latent structure is continuous (Bauer 2007; Bauer and Curran 2003; McGrath and Walters 2012); in fact, a review of 59 group-based modeling studies on externalizing problems (Van Dulmen et al. 2009) failed to identify a single one-component outcome.

The data used in the current study can be viewed as fairly representative of prior research on developmental trajectories to the extent that GMM and LCGA, using the LMR stopping rule identified two (aggression, noncompliance) and three (callousness) groups, with features similar to those obained in previous research (Broidy et al. 2003; van Dulmen et al. 2009). The two-class solution is not, in fact, incongruent with Moffitt's (1993) theory, given the age at which the follow-up ended: i.e., the AL pattern does not normally start until around age 14. Hence, dividing the distribution into LCP and non-LCP cases makes perfect sense from the standpoint of Moffitt's theory. If these trajectories reflected meaningful and replicable categories of antisocial behavior then there should have been evidence of at least one taxonic boundary when the three antisocial indicators were compared across and within time periods. The fact that there were no signs of categorical latent structure in taxometric analyses of antisocial behavior across a time series spanning 8 years, from age 6 to age 14, or within any of the

Table 3 Descriptive statistics, indicator validities, and mean inter-indicator correlations for the three indicators from the longitudinal and cross-sectional analyses of the total sample

Indicators	Range	*M*	SD	Skew	Kurtosis	*d*	*r*(tot)	*r*(tax)	*r*(comp)
Summed indicators						2.29(2.58)	0.58	0.16(0.07)	0.20(0.31)
Callousness	0–19	4.09	3.36	0.77	0.15	2.36(2.27)			
Aggression	0–17	1.79	2.50	1.93	4.23	2.34(2.93)			
Noncompliance	0–16	1.78	2.63	1.93	4.02	2.18(2.54)			
Age 6 indicators						1.94(2.24)	0.37	−0.02(−0.02)	−0.01(0.05)
Callousness	0–4	0.87	0.89	0.78	0.09	2.04(1.76)			
Aggression	0–4	0.36	0.67	1.97	3.86	2.15(2.37)			
Noncompliance	0–4	0.25	0.60	2.82	8.96	1.64(2.58)			
Age 8 indicators						1.98(2.39)	0.41	0.01(0.02)	−0.00(0.12)
Callousness	0–4	0.86	0.91	0.88	0.21	2.16(1.85)			
Aggression	0–4	0.38	0.70	2.15	5.17	2.08(2.74)			
Noncompliance	0–4	0.30	0.65	2.36	5.95	1.70(2.58)			
Age 10 indicators						2.01(2.15)	0.39	−0.09(−0.09)	−0.01(0.07)
Callousness	0–4	0.79	0.89	0.88	0.04	2.07(1.73)			
Aggression	0–4	0.33	0.64	2.04	4.03	2.08(2.23)			
Noncompliance	0–4	0.34	0.72	2.24	5.01	1.88(2.48)			
Age 12 indicators						2.07(2.36)	0.43	−0.02(−0.06)	0.01(0.10)
Callousness	0–4	0.77	0.91	1.01	0.37	2.15(2.06)			
Aggression	0–4	0.34	0.66	2.16	5.17	1.98(2.33)			
Noncompliance	0–4	0.40	0.79	2.19	4.83	2.09(2.68)			
Age 14 indicators						2.40(2.48)	0.50	0.04(0.04)	0.10(0.16)
Callousness	0–4	0.79	0.99	1.10	0.44	2.42(2.23)			
Aggression	0–4	0.38	0.72	2.18	5.10	2.22(2.48)			
Noncompliance	0–4	0.49	0.90	1.83	2.84	2.55(2.75)			

First figure is based on the GMM-derived taxon base rate of 0.242 and the second figure (in parentheses) is based on MAMBAC/MAXCOV-derived taxon base rates

Range range of scores, *M* mean, *SD* standard deviation, *d* Cohen's *d*, *r*(tot) mean inter-indicator correlation for total sample, *r*(tax) mean inter-indicator correlation for the putative taxon, *r*(comp) mean inter-indicator correlation for the putative complement

individual time periods, suggests that a meaningful discontinuity did not exist in the distribution of antisocial scores examined in this study.

Unlike taxometrics, GMM and LCGA identified a small cluster of LCP or high antisocial profiles and several additional categories depending on whether the LMR or AIC/BIC stopping rule was used. Owing to the fact that GMM and LCGA account for non-normality in a continuous or admixed distribution by creating two or more normal component distributions, it was felt that these results should be verified against a procedure based on a different set of assumptions. When the alternate procedure (i.e., taxometrics) was applied to the data it failed to identify an underlying discontinuity in the distribution of antisocial ratings. It should be noted, however, that the raw data used in the mixture modeling and taxometric analyses, while related, were not exactly the same. Whereas GMM and LCGA were used to analyze and estimate growth parameters (slopes and intercepts) in the three antisocial indicators (callousness, aggression, noncompliance) across five age periods, the taxometric method was used to analyze the variance/covariance matrices of scores on the three antisocial indicators (callousness, aggression, noncompliance) summed across the five age periods. Differences in how the three antisocial indicators were analyzed taxometrically and with growth mixture modeling could therefore account for some of the differences observed in outcome between the two procedures.

The use of indicators from a single source (behavioral ratings) and the utilization of summed scores are two potential limitations of this study. Basing one's indicators on a single data source leaves one vulnerable to charges of mono-operation bias (Shadish et al. 2002) and suggests the need for replication of one's findings in a new sample with alternate measures, such as teacher- or self-reports of antisocial behavior. Research, in fact, indicates that rater agreement is only modest to moderate and influenced by a number of factors (Gresham et al. 2010). Beyond the contextual and attitudinal differences that exist between teacher and parent raters, teacher

Table 4 Correlation matrix for the three indicators across the five time periods

	C6	A6	N6	C8	A8	N8	C10	A10	N10	C12	A12	N12	C14	A14	N14
Callousness-6yo		0.50	0.35	0.47	0.38	0.28	0.42	0.32	0.29	0.39	0.30	0.29	0.36	0.29	0.32
Aggression-6yo	0.38		0.36	0.40	0.53	0.31	0.36	0.44	0.27	0.30	0.37	0.25	0.30	0.36	0.27
Noncompliance-6yo	0.27	0.30		0.28	0.22	0.44	0.24	0.18	0.29	0.24	0.20	0.28	0.21	0.18	0.29
Callousness-8yo	0.38	0.34	0.28		0.48	0.37	0.52	0.32	0.31	0.41	0.29	0.26	0.42	0.31	0.30
Aggression-8yo	0.25	0.40	0.19	0.49		0.39	0.36	0.51	0.34	0.37	0.47	0.31	0.37	0.44	0.29
Noncompliance-8yo	0.19	0.13	0.32	0.32	0.37		0.28	0.25	0.52	0.31	0.27	0.47	0.32	0.30	0.41
Callousness-10yo	0.36	0.25	0.19	0.47	0.26	0.21		0.45	0.34	0.43	0.30	0.26	0.39	0.30	0.31
Aggression-10yo	0.26	0.37	0.16	0.36	0.44	0.20	0.41		0.44	0.36	0.51	0.32	0.34	0.48	0.34
Noncompliance-10yo	0.16	0.11	0.19	0.20	0.13	0.34	0.33	0.28		0.33	0.31	0.52	0.35	0.30	0.44
Callousness-12yo	0.34	0.26	0.16	0.38	0.25	0.17	0.47	0.32	0.23		0.47	0.44	0.53	0.42	0.36
Aggression-12yo	0.22	0.33	0.12	0.26	0.37	0.16	0.27	0.47	0.18	0.42		0.42	0.37	0.53	0.35
Noncompliance-12yo	0.20	0.19	0.22	0.21	0.22	0.29	0.24	0.32	0.35	0.37	0.42		0.38	0.37	0.56
Callousness-14yo	0.28	0.21	0.15	0.37	0.22	0.16	0.40	0.27	0.23	0.47	0.35	0.32		0.56	0.49
Aggression-14yo	0.19	0.24	0.10	0.25	0.30	0.14	0.26	0.35	0.20	0.32	0.47	0.30	0.53		0.43
Noncompliance-14yo	0.16	0.11	0.14	0.21	0.15	0.23	0.24	0.25	0.31	0.26	0.31	0.39	0.48	0.47	

Correlations above the diagonal are for males (n=843) and correlations below the diagonal are for females (n=865); all correlations significant at p<0.01

ratings appear to be less influenced by anxiety and depression and the current family situation than parental ratings (De Los Reyes and Kazdin 2005). The use of longitudinal scores summed across five developmental periods (ages 6–14) could also be considered a limitation of this study. Summed scores may account for longitudinal differences in magnitude between patterns but cannot explain potentially important curvilinear relationships within a time series. Nevertheless, the

Table 5 Results of taxometric analyses for the total sample and male and female subsamples

Ages	Taxon base rate estimate			Comparison Curve Fit Index (CCFI)			
	MAMBAC	MAXCOV	Mean	MAMBAC	MAXCOV	L-Mode	Mean
Total sample (N=1,708)							
Sum	0.173	0.130	0.152	0.254	0.166	0.346	0.255
Age 6	0.153	0.110	0.132	0.437	0.363	0.387	0.396
Age 8	0.115	0.118	0.116	0.430	0.197	0.239	0.289
Age 10	0.169	0.173	0.171	0.348	0.385	0.330	0.354
Age 12	0.130	0.149	0.140	0.254	0.227	0.373	0.285
Age 14	0.160	0.126	0.143	0.319	0.110	0.248	0.226
Male subsample (n=843)							
Sum	0.229	0.163	0.196	0.274	0.256	0.337	0.289
Age 6	0.147	0.125	0.136	0.456	0.270	0.355	0.360
Age 8	0.230	0.156	0.193	0.398	0.242	0.249	0.296
Age 10	0.201	0.224	0.212	0.530	0.296	0.336	0.387
Age 12	0.200	0.226	0.210	0.385	0.332	0.332	0.350
Age 14	0.209	0.200	0.204	0.423	0.286	0.324	0.344
Female subsample (n=865)							
Sum	0.181	0.112	0.146	0.440	0.298	0.408	0.382
Age 6	0.074	0.106	0.090	0.553	0.367	0.377	0.432
Age 8	0.074	0.090	0.082	0.505	0.424	0.438	0.456
Age 10	0.124	0.184	0.154	0.483	0.476	0.346	0.435
Age 12	0.165	0.114	0.140	0.468	0.272	0.344	0.361
Age 14	0.137	0.091	0.114	0.482	0.286	0.284	0.351

MAMBAC mean above minus below a cut; *MAXCOV* maximum covariance; *L-Mode* latent mode; *Mean* mean of MAMBAC, MAXCOV, and L-Mode

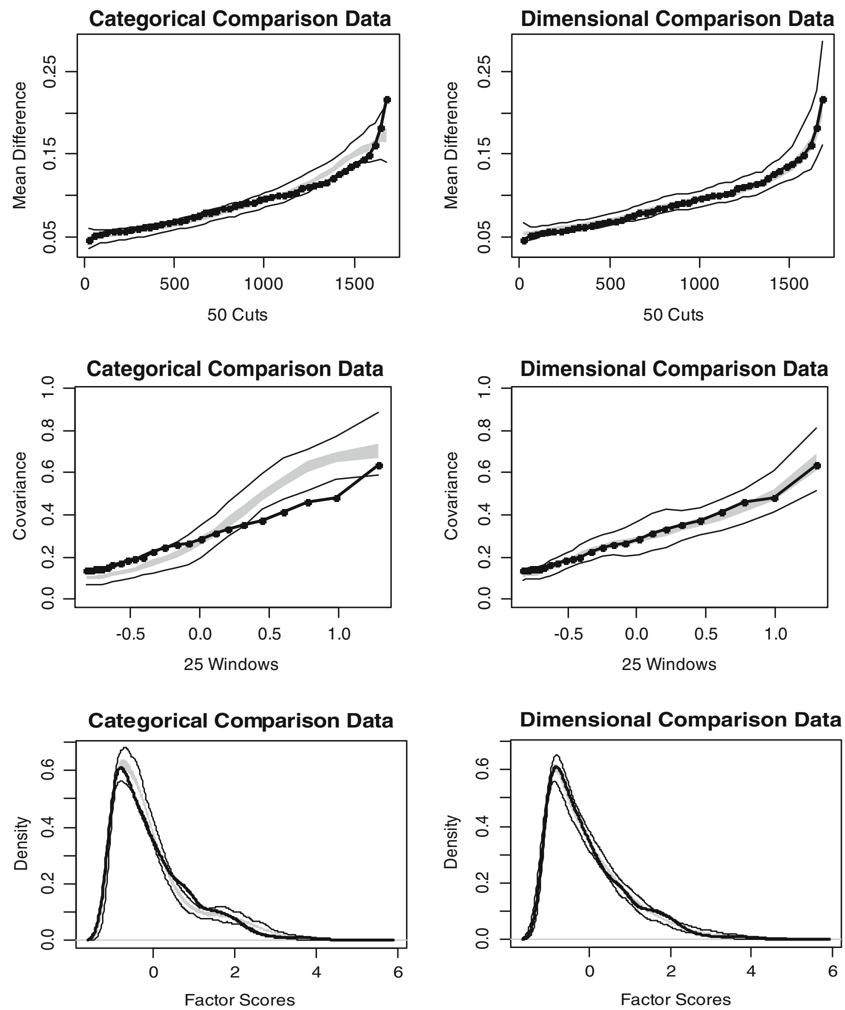


Fig. 3 Taxometric results for the analysis of indicators representing the summed (longitudinal) data in the total sample. Categorical comparison data are based on an average MAMBAC-MAXCOV putative taxon base rate of 0.152. Results are shown for MAMBAC (mean above minus below a cut; *top panel*), MAXCOV (maximum covariance; *middle panel*), and L-Mode (latent mode; *bottom panel*) procedures. In each graph, the *dark line* shows results for the empirical data, *lighter lines* connect the minimum and maximum values at each data point, and the *gray area* represents the middle 50 % of values. For each procedure, the results for the empirical data resemble those for the dimensional (continuous) comparison data more closely than those for the categorical comparison data

GMM and LCGA procedures identified between one and seven boundaries in the current data and none were detected by the taxometric method despite the fact the trajectories identified by GMM and LCGA differed primarily in magnitude.

Just as growth mixture modeling has limitations (Bauer 2007), so too does the taxometric method. First, the indicators in a taxometric analysis should correlate at least moderately (mean $r \geq 0.30$) in the full sample but less than moderately (mean $r < 0.30$) in the putative taxon and complement samples (Meehl 1995). Although analyses conducted using GMM-generated base rates produced results congruent with Meehl's (1995) recommendations, analyses using base rates obtained from the taxometric analyses disclosed correlations within the complement group that were slightly above Meehl's (1995) recommended threshold. This was neither unexpected nor particularly problematic, however, given the low base rate (0.15) and a moderately high mean inter-item correlation in the full sample (0.58). Second, the indicators must be capable of distinguishing between the putative taxon and complement groups at a high level of differentiation ($d \geq 1.25$). This requirement will make certain samples inappropriate for taxometrics, although in the current study it did not present a problem in analyses performed using either the GMM- or taxometric-generated base rates. Third, the taxometric method has traditionally only been employed with cross-sectional data. The present study sought to increase the relevance of taxometrics to developmental research by examining its performance with longitudinal data. GMM and LCGA indicated the presence of two to eight distinct trajectories and a fairly consistent boundary across procedures, indicators, and stopping rules when distinguishing between low and high/moderate elevation in the antisocial indicators. Since taxometrics is designed for only two outcomes (categorical, continuous) it could be argued that as the number of legitimate categories rise, taxometrics become less effective in determining latent structure. Monte Carlo research, however, indicates that the taxometric method can detect categorical latent structure in samples containing at least five classes (Frazier et al. 2011; McGrath and Walters 2012; Walters et al. 2010).

Because the use of scores summed over several developmental periods in a taxometric investigation is a novel application, ancillary simulations were conducted to assess

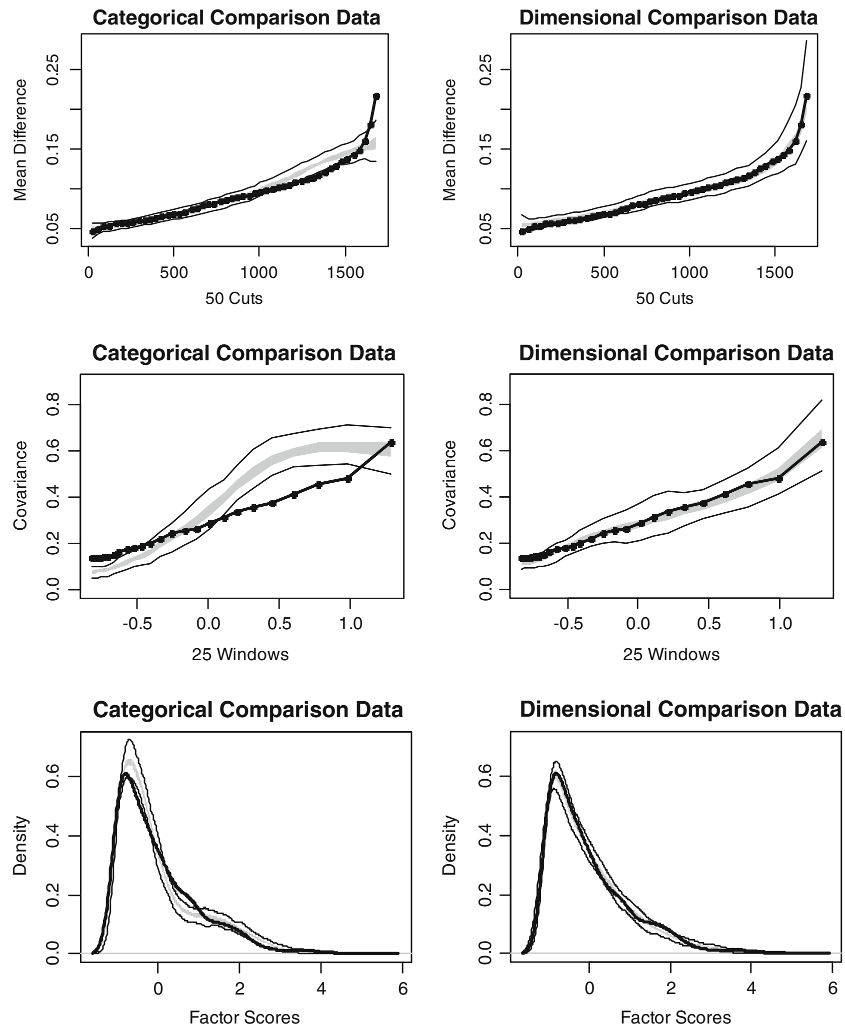


Fig. 4 Taxometric results for the analysis of indicators representing the summed (longitudinal) data in the total sample. Categorical comparison data are based on GMM results averaged across the three indicators using both the LMR and AIC/BIC stopping rules to create a putative taxon base rate of 0.242 (high/moderate vs. low elevation). Results are shown for MAMBAC (mean above minus below a cut; *top panel*), MAXCOV (maximum covariance; *middle panel*), and L-Mode (latent mode; *bottom panel*) procedures. In each graph, the *dark line* shows results for the empirical data, *lighter lines* connect the minimum and maximum values at each data point, and the *gray area* represents the middle 50 % of values. For each procedure, the results for the empirical data resemble those for the dimensional (continuous) comparison data more closely than those for the categorical comparison data

the method's performance with longitudinal data similar to the data upon which the current study was based. Constructing artificial categorical data based on the parameters obtained from the growth mixture modeling results, it was possible to examine taxometric results for longitudinal data modeled after the data used in the current study. Curves based on the artificial data were markedly different from curves produced when the empirical data were analyzed, suggesting that the current data were probably appropriate for taxometric analysis. Before a definitive answer can be found for the larger question of whether the taxometric method is applicable to many forms of longitudinal data, however, a more comprehensive Monte Carlo analysis must be performed. In planning future research on the utility of taxometric procedures with longitudinal data, we might want to take note of one additional advantage the taxometric method has over growth mixture modeling and related procedures. By allowing for both longitudinal analysis of summed data and cross-sectional analysis of the individual points in a time series the taxometric method expands the number of analyses and with this the number of opportunities for consistency testing, a hallmark of Meehl's (1995, 2004) taxometric approach.

Nagin and Tremblay (2005) contend that detractors have misinterpreted the results of mixture modeling studies by incorrectly assuming individual members of a trajectory class actually belong to the trajectory, the number of trajectory groups is immutable, and individual trajectory group members follow group-level trajectories in lock step. They go on to argue that above all else, multiple trajectory models have heuristic value in that they can serve to clarify complex and contradictory information. This may be true but it is also important to validate these preliminary findings against alternative statistical procedures given the fact that mixture modeling will create components out of non-normality regardless of whether the underlying structure is categorical or continuous. The current study attempted to do just this by subjecting reconstructed trait scales from the antisocial score of the BPI to taxometric analysis. The results indicated that from age 6 to 14 there was no evidence of discontinuity in the latent structure of youthful antisocial behavior. Future research is required, however, to determine whether this conclusion holds when older adolescents are examined and the AL pattern is prominent. Latent structure is important and this initial application of the taxometric procedure using longitudinal data suggests that the development of antisocial behavior, at least with respect to maternal ratings of children in the NLSY-C sample, falls along a continuum rather than into distinct categories. This is a finding with potentially important implications for ongoing research on developmental patterns of antisocial behavior.

References

Baker, P., Keck, C. K., Mott, F. L., & Quinlan, S. V. (1993). *NLSY child handbook: A guide to the 1986–1990 NLSY child data (rev ed.)*. Columbus: The Ohio State University, Center for Human Resource Research.

Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology, 28*, 97–104.

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research, 42*, 757–786.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychological Methods, 8*, 338–363.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, R., Dodge, K., et al. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: a six-site, cross-national study. *Developmental Psychology, 39*, 222–245.

Burt, S. A., Donnellan, M. B., Iacono, W. G., & McGue, M. (2011). Age-of-onset or behavioral sub-types? A prospective comparison of two approaches to characterizing the heterogeneity with antisocial behavior. *Journal of Abnormal Psychology, 39*, 633–644.

Bushway, S. D., Thornberry, T. P., & Krohn, M. D. (2003). Desistance as a developmental process: a comparison of static and dynamic approaches. *Journal of Quantitative Criminology, 19*, 129–153.

Center for Human Resource Research (2009). *NLSY79 user's guide*. The Ohio State University, CHRR NLS User Services.

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.

Cronbach, L., & Furby, L. (1970). How we should measure 'change': or should we? *Psychological Bulletin, 74*, 68–80.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483–509.

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.

Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M. J., & Kettler, R. (2010). Cross-informant agreement for ratings for social skill and problem behavior ratings: an investigation of the Social Skills Improvement System-Rating Scales. *Psychological Assessment, 22*, 157–166.

Grove, W. M., & Meehl, P. E. (1993). Simple regression based procedures for taxometric investigation. *Psychological Reports, 73*, 707–737.

Harzing, A. W. (2007). *Publish or perish*. (http://www.harzing.com/pop.htm). Accessed 28 July 2012.

Haslam, N., & Cleland, C. (2002). Taxometric analysis of fuzzy categories: a Monte Carlo study. *Psychological Reports, 90*, 401–404.

Hoeve, M., Blokland, A., Dubas, J. S., Loeber, R., Gerris, J. R. M., & Van der Laan, P. H. (2008). Trajectories of delinquency and parenting styles. *Journal of Abnormal Child Psychology, 36*, 223–235.

Lacourse, E., Cote, S., Nagin, D. S., Vitaro, F., Brendgen, M., & Tremblay, R. E. (2002). A longitudinal–experimental approach to testing theories of antisocial behavior. *Development and Psychopathology, 14*, 909–924.

Lahey, B. B., Van Hulle, C. A., Waldman, I. D., Rodgers, J. L., D'Onofrio, B. M., & Pedlow, S. (2006). Testing descriptive hypotheses regarding sex differences in the development of conduct problems and delinquency. *Journal of Abnormal Child Psychology, 34*, 737–755.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767–778.

McGrath, R. E., & Walters, G. D. (2012). Taxometric analysis as a general strategy for distinguishing categorical from dimensional latent structure. *Psychological Methods, 17*, 284–293.

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

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

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

Meehl, P. E. (2004). What's in a taxon? *Journal of Abnormal Psychology, 113*, 39–43.

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.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review, 100*, 674–701.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park: Sage Publications.

Muthén, B., & Muthén, L. (1998–2007). *Mplus user's guide* (5th ed.). Los Angeles: Muthén & Muthén.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semi-parametric group-based approach. *Psychological Methods, 4*, 139–157.

Nagin, D. S., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development, 70*, 1181–1196.

Nagin, D. S., & Tremblay, R. E. (2005). Developmental trajectory groups: fact or a useful statistical fiction? *Criminology, 43*, 873–904.

Nagin, D. S., Farrington, D. P., & Moffitt, T. E. (1995). Life-course trajectories of different types of offenders. *Criminology, 33*, 111–139.

Peterson, J. L., & Zill, N. (1986). Marital disruption, parent–child relationships, and behavioral problems in children. *Journal of Marriage and Family, 48*, 295–307.

R Development Core Team (2005). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Ruscio, J. (2007). Taxometric analysis: an empirically-grounded approach to implementing the method. *Criminal Justice and Behavior, 34*, 1588–1622.

Ruscio, J. (2009). Assigning cases to groups using taxometric results: an empirical comparison of classification techniques. *Assessment, 16*, 55–70.

Ruscio, J. (2011). *Taxometric programs for the R computing environment: User's manual*. (http://www.tcnj.edu/~ruscio/taxometrics.html). Accessed 3 November 2011.

Ruscio, J., & Kaczetow, W. (2008). Simulating multivariate nonnormal data using an iterative algorithm. *Multivariate Behavioral Research, 43*, 355–381.

Ruscio, J., & Marcus, D. K. (2007). Detecting small taxa using simulated comparison data: a reanalysis of Beach, Amir, and Bau's (2005) data. *Psychological Assessment, 19*, 241–246.

Ruscio, J., & Ruscio, A. M. (2004). A conceptual and methodological checklist for conducting a taxometric investigation. *Behavior Therapy, 35*, 403–447.

Ruscio, J., & Walters, G. D. (2009). Using comparison data to differentiate categorical and dimensional data by examining factor score distributions: resolving the mode problem. *Psychological Assessment, 21*, 578–594.

Ruscio, J., & Walters, G. D. (2011). Differentiating categorical and dimensional data with taxometric analysis: are two variables better than none? *Psychological Assessment, 23*, 287–299.

Ruscio, J., Haslam, N., & Ruscio, A. M. (2006). *Introduction to the taxometric method: A practical guide*. Mahwah: Lawrence Erlbaum.

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.

Ruscio, J., Walters, G. D., Marcus, D. K., & Kaczetow, W. (2010). Comparing the relative fit of categorical and dimensional latent variable models using consistency tests. *Psychological Assessment, 22*, 5–21.

Sampson, R. J., & Laub, J. H. (2005). Seductions of method: rejoinder to Nagin and Tremblay's "developmental trajectory groups: fact or fiction?". *Criminology, 43*, 905–913.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Mifflin.

Van Dulmen, M., Goncy, E., Vest, A., & Flannery, D. (2009). Group-based trajectory modeling of externalizing behavior problems from childhood through adulthood: Exploring discrepancies in the empirical findings. In J. Savage (Ed.), *The development of persistent criminality* (pp. 288–314). New York: Oxford University Press.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks: 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.

Walters, G. D. (2012). Taxometrics and criminal justice: assessing the latent structure of crime-related constructs. *Journal of Criminal Justice, 40*, 10–20.

Walters, G. D., & Ruscio, J. (2009). To sum or not to sum: taxometric analysis with ordered categorical assessment items. *Psychological Assessment, 21*, 99–111.

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.

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.