The dimensional nature of externalizing behaviors in adolescence:

Evidence from a direct comparison of categorical, dimensional, and hybrid models

Kate E. Walton

St. John's University

Johan Ormel

University of Groningen

Robert F. Krueger

University of Minnesota

Correspondence should be sent to:

Kate E. Walton

Department of Psychology

St. John's University

Marillac Hall sb36

8000 Utopia Parkway

Jamaica, NY 11439

waltonk@stjohns.edu

Abstract

Researchers have recognized the importance of developing an accurate classification system for externalizing disorders, though much of this work has been framed by a priori preferences for categorical vs. dimensional constructs. Newer statistical technologies now allow categorical and dimensional models of psychopathology to be compared empirically. In this study, we directly compared the fit of categorical and dimensional models of externalizing behaviors in a large and representative community sample of adolescents at two time points separated by nearly 2.5 years (N = 2027; mean age at Time 1 = 11.09 years; 50.8% female). Delinquent and aggressive behaviors were assessed with child and parent Child Behavior Checklist reports. Latent trait, latent class, and factor mixture models were fit to the data, and at both time points, the latent trait model provided the best fit to the data. The item parameters were inspected and interpreted, and it was determined that the items were differentially sensitive across all regions of the dimension. We conclude that classification models can be based on empirical evidence rather than a priori preferences, and while current classification systems conceptualize externalizing problems in terms of discrete groups, they can be better conceptualized as dimensions.

The dimensional nature of externalizing behaviors in adolescence: Evidence from a direct comparison of categorical, dimensional, and hybrid models

Externalizing behaviors in adolescents, such as aggression and delinquency, are highly problematic for society. In 2008, nearly 1.2 million American children under the age of 18 were arrested for crimes ranging from curfew and loitering violations, to rape and murder (U.S. Department of Justice, Federal Bureau of Investigation, 2009). The majority of these offenses were for property crime and larceny or theft, and there were also an alarming number of arrests for drug and alcohol-related violations and assault. Self-reports also point to high rates of delinquent and aggressive behavior. According to a 2007 study conducted by the Centers for Disease Control and Prevention, nearly 45% of American high school students have used alcohol in the past month, 19.7% have used marijuana in the past month, and 35.5% have been in a physical fight in the past year. Similar prevalence rates are observed in other developed countries. For example, in The Netherlands, 50.5% of high school students used alcohol in the past month, 13.5% used marijuana in the past year, and 29% were in a physical altercation in the past six months (Monshouwer et al., 2008). These statistics are troubling because the effects of externalizing behavior go well beyond the immediate adverse consequences for the public and the individual. For example, longitudinal research shows that adolescent delinquency leads to decreased educational and occupational attainment in adulthood (e.g., Laub & Sampson, 1994; Tanner, Davis, & O'Grady, 1999), and low attainment may mediate the relationship between adolescent delinquency and depression in young adulthood (Siennick, 2007).

Given the prevalence of externalizing behaviors and their consequences in both adolescence and adulthood, researchers have recognized the importance of understanding the nature of these behaviors and developing an accurate classification system. Many have conceptualized externalizing problems as discrete conditions. Most major psychiatric classification systems treat various types of externalizing problems as categorically distinct from one another, other types of psychopathology, and normal functioning (e.g., American Psychiatric Association, 1994). Theoretically, such an account is plausible if the underlying etiology of these conditions is highly discrete (e.g., a specific genetic polymorphism has a major and indelible effect, leading to discrete classes of those who show specific externalizing behaviors and those who do not). In contrast, others have argued that externalizing behaviors may be better conceptualized in dimensional terms (e.g., Dick et al., 2008; Kendler, Prescott, Myers, & Neale, 2003; Krueger, Markon, Patrick, & Iacono, 2005). Theoretically, such an account is plausible if the underlying etiology of these conditions is multifaceted, encompassing a variety of genetic polymorphisms and environmental elicitors. It is also possible that the latent structure of externalizing psychopathology contains both continuous and discrete elements. For example, there may be continuous variation in a "normal" range, then a threshold beyond which manifest externalizing problems become dramatically more pronounced, with continuous variation in both the "normal" (below the threshold) and "pathological" (above the threshold) classes. Numerous genetic polymorphisms and environmental elicitors could combine to create a continuum of risk, but there may also be a threshold, beyond which the social problems linked to externalizing accelerate. This could create a latent discontinuity, as well as continuous variation both above and below the point of discontinuity.

Much work on classification has been framed by a priori preferences for categorical vs. dimensional constructs (e.g., see Widiger & Samuel, 2005, for a discussion). However, newer statistical technologies now allow categorical and dimensional models of psychopathology to be compared empirically (Lubke & Neale, 2006; Markon & Krueger, 2006; Muthén, 2006). In this study, we used statistical modeling to shed light on the empirical structure of externalizing behaviors (specifically, delinquent and aggressive behaviors as conceptualized by Achenbach [1991a]) in early adolescence. To be precise, we directly compared the fit of categorical (i.e., latent class), dimensional (i.e., latent trait), and hybrid (i.e., factor mixture) models of externalizing behaviors in a large and representative community sample of adolescents. We now turn to consider how this direct comparison extends our current understanding of the nature of adolescent externalizing problems.

Structural Modeling of Externalizing Psychopathology in Adolescence

Two basic types of structural models are particularly relevant to understanding the nature of adolescent externalizing behaviors: latent class and latent trait models. Latent class models model patterns of symptoms in terms of a specific number of mutually exclusive groups where group membership is associated with specific probabilities of certain symptoms being present (McCutcheon, 1987). In contrast, latent trait models (also referred to as item response models) model symptoms in terms of their relationship with an underlying continuum (Hambleton, Swaminathan, & Rogers, 1991). A third type of model is a hybrid model, providing a categorical and continuous representation of latent variables. Factor mixture models categorize people into discrete classes, but unlike latent class models, there is within class heterogeneity (Muthén, 2006).

A number of latent structure modeling studies of delinquent and aggressive behaviors in adolescence have been reported. In a sample of adolescent males, Brownfield and Sorenson (1987) modeled indicators of delinquency involving theft, vandalism, and assault. Three qualitatively distinct classes emerged. In a study of adolescent females, Odgers et al. (2007) fit a latent class model to self-reports of delinquent and violent behavior, and three classes emerged. Finally, latent class analyses of conduct disorder symptoms suggest a number of distinct subtypes of individuals (Eaves et al., 1993; Nock, Kazdin, Hiripi, & Kessler, 2006).

In contrast, latent trait modeling of conduct disorder symptoms has produced evidence for a continuous distribution, countering the idea that there are qualitatively distinct groups of delinquents (Gelhorn et al., 2009). The same has been found for latent trait modeling of general delinquency indicators (Van den Oord, Pickles, & Waldman, 2003). Finally, latent trait modeling of adolescent psychopathy (a personality style associated with callous affect, interpersonal manipulation, impulsivity, and delinquent behavior) has also provided support for a dimensional structure (Schrum & Salekin, 2006).

In sum, findings from latent class and latent trait analyses in the adolescent literature have provided contradictory findings, some supporting a categorical structure and some supporting a continuous dimension of externalizing behaviors, based on the type of model fit in the various studies. However, no published studies that we are aware of have fit these distinct latent structure models in a single study of adolescents in order to compare their fit (and none have employed factor mixture modeling). The value of testing the relative fit of distinct models has been demonstrated, as it allows one to correctly distinguish between discrete and continuous latent variables (Lubke & Neale, 2006; Markon & Krueger, 2006; Muthén, 2006). For example, Krueger et al. (2005), and Markon and Krueger (2005) directly compared the relative fit of latent class and latent trait models to externalizing behaviors in adulthood. In both studies, the authors concluded that a latent trait model with a single underlying continuum was the best fitting model.

The Current Study

Latent structure modeling of externalizing behaviors in adolescence has not lead to a definitive conclusion regarding their latent structure. However, few studies have incorporated both latent class and latent trait models in a single study of externalizing, and only two (Krueger et al., 2005; Markon & Krueger, 2005), both of which focused solely on the structure of externalizing in adulthood, have directly compared the fit of the two models. No studies we are aware of have directly compared the fit of categorical, continuous, and hybrid latent variable models in an adolescent sample. Furthermore, no studies have assessed the relative fit of these models over time. An understanding of the structural consistency over time would frame our conceptualization of continuity at the phenotypic level. If individuals fall into distinct classes, the stability of class membership over time would be relevant (e.g., Lanza & Collins, 2008), whereas if individuals are ordered along a common continuum, their differential stability over time would be relevant.

In this study, we compared the fit of latent trait, latent class, and factor mixture models to aggression and delinquency indicators in a longitudinal study of a large, population-based sample of adolescents. The participants were assessed at two time points (ages 11 and 13) via parent and self-report. To our knowledge, this is the first study to directly compare the fit of continuous, categorical, and hybrid models of externalizing behaviors in an adolescent sample, and is the first to assess the fit at two time points and to discern the implications this has for the conceptualization of continuity and change over time.

Method

Participants

Participants were part of an ongoing prospective cohort study of Dutch adolescents, the Tracking Adolescents' Individual Lives Survey (TRAILS). The primary objective of TRAILS is to investigate the development of mental health from preadolescence into adulthood. The baseline survey and all subsequent waves were approved by the national ethical committee for human research, 'Centrale Commissie Mensgebonden Onderzoek', and parents signed an informed consent form at the beginning of the study. All children born in 1991 in the three largest cities and two rural areas in the north of The Netherlands were invited for participation. A more detailed description of TRAILS, including sampling methods and study design, can be found in De Winter et al. (2005) and Huisman et al. (2008). The first wave of assessment (T_1) ran from March 2001 to July 2002. Of all eligible children, 76.0% (N = 2230; mean age = 11.09, SD = .55; 50.8% female) were enrolled in the study (i.e., both child and parent agreed to participate). The second wave of assessment (T_2) occurred between September 2003 and December 2004. More than 96% of the original sample participated (N = 2149). At this time, the mean age of the children was 13.55 years (SD = .54), and 51.2% were female. For this study, cases were eliminated if they were missing all or most (i.e., completed only one or two items) of the self or parent report. As a result, the final sample sizes were 2027 at T_1 and 1890 at T_2.

To enable a comprehensive analysis of non-response bias, information on mental health determinants and outcomes was obtained from the teachers of responders and non-responders

(De Winter et al., 2005). Although responders and non-responders did not differ (χ^2, $p > .05$) with respect to the prevalence of teacher-rated problem behaviors at T_1 (aggression: 18.1% vs. 18.0%, r_ϕ =.00; rule breaking: 7.1% vs. 7.8%, $r_\phi = .01$) and single parent family status (15.3% vs. 16.6%, $r_\phi = .01$), non-responders were more often boys (55.7% vs. 49.2%, $r_\phi = .06$), and the proportion with low parental education level was higher in the non-responders (48.7% vs. 37.7%, $r_\phi = .08$). The need for additional help due to learning difficulties was also higher in non-responders (28.8% vs. 21.1%, $r_\phi = .06$). No differences between responders and non-responders were found regarding associations between sociodemographic variables (gender, age, SES) and problem behaviors at T_1 (De Winter et al., 2005). At T_2, no differences were found between responders and non-responders, but low family SES children were overrepresented in the non-responders (47.2% vs. 24.5%, $r_\phi = .09$).

We also examined differences between those who participated at both time points vs. those who participated only at T_1. Parent-reported aggression ($t = 4.05$, $p < .05$; $d = .27$) and rule breaking behavior ($t = 5.34$, $p < .05$; $d = .36$) were higher for those who did not participate at both time points, as was self-reported aggression ($t = 2.09$, $p < .05$; $d = .14$). There was no difference on self-reports of rule breaking behavior ($t = 1.53$, $p > .05$; $d = .10$).

Measures

At both time points, parents completed the Dutch version of the Child Behavior Checklist (CBCL; Achenbach, 1991a; Verhulst, Van der Ende, & Koot, 1996). The CBCL was designed for parents to assess behavioral and emotional problems of their children. It consists of 120 items scored on a three-point scale where *0* indicates the item is not true, *1* indicates the item is somewhat/sometimes true, and *2* indicates the item is very/often true. For the purpose of this paper, we used two scales of the CBCL that focus on externalizing problems (rule breaking behavior and aggression). The children also completed a self-report version of the CBCL, the Youth Self-Report (Achenbach, 1991b). There were 15 overlapping (self-report to parent-report) items for rule breaking and 17 items for aggression. This set of items taps many key components of externalizing. Included are several personality-based items (e.g., lack of feelings of guilt, moodiness, and irritability), as well as items tapping specific behaviors ranging from substance use (nicotine, alcohol, and illicit drugs) to relational and physical aggression.

Ratings from parents and children were combined by using the higher score for each item (Frick, Cornell, Barry, Bodin, & Dane, 2003; Piacentini, Cohen, & Cohen, 1992). We deemed it important to consider both child and parent reports as each offers a unique perspective. In fact, there is often a lack of agreement between self- and observer-reports for both child (De Los Reyes & Kazdin, 2005) and adult psychopathology with mean self-other correlations not exceeding .44 for externalizing problems (Achenbach, Krukowski, Dumenci, & Ivanova, 2005). Children are more privy to their thoughts and feelings (e.g., feeling guilty) while parents may provide more accurate assessments of certain behaviors (e.g., being overly loud). For example, children report significantly less disruptive behavior than their parents and teachers (Loeber, Green, Lahey, & Stouthamer-Loeber, 1991). The primary motive for using the combined approach was to yield more realistic prevalence rates, as children as well as parents might be inclined to underreport various behaviors. Therefore, considering an item to be “very/often true” only when both informants selected that response option did not seem reasonable. In the event that a case was missing the child or parent report, the one existing report was used (on average, this occurred for less than 1% of the cases at T_1 and T_2). Items were dichotomized so *0* indicates the item is not true, and *1* indicates the item is somewhat/sometimes or very/often true. This was done because the highest category was rarely endorsed for a number of items, and low frequencies in any one response category can lead to problems in the estimation of the parameters of statistical models, such as those used in this study.

Statistical Modeling

In order to determine whether facets of externalizing behaviors are continuous or categorical in nature, we fit three types of models (latent trait model, latent class, and factor mixture models), and compared their relative fit. Latent trait models (Hambleton et al., 1991) describe the probability of a discrete outcome (e.g., endorsing an item) as a function of person parameters (individuals' levels on a continuous latent trait, referred to as theta) and item parameters. Specifically, we fit a two-parameter logistic model (2-PL). The first item parameter is a discrimination parameter, which indicates the ability to differentiate among individuals at varying levels of the latent trait. Items that are not highly relevant to the latent trait will have discrimination values approaching 0. The second is a location parameter (also referred to as a difficulty or threshold parameter), which indicates the level of the latent trait necessary to have a probability of .5 to endorse an item (the location parameter is on the same *z*-score metric as theta). For example, items that are not likely to be endorsed because of their extreme nature have high location parameters.

The objective with latent class modeling is to categorize individuals into discrete classes using a set of indicators (McCutcheon, 1987). Similar to latent trait models, two types of parameters are estimated with latent class models, in this case, item parameters and class probability parameters. With categorical indicators, the item parameters are conditional item probabilities, which are specific to each class and indicate the likelihood of an individual in that class to endorse a given item. The class probability parameters specify the proportion of the population in each class.

Factor mixture models are hybrid models in that they simultaneously provide categorical and continuous representation of underlying latent variables. Individuals are categorized into discrete classes, but within each class, a continuous latent factor accounts for differences in the severity of the disorder. In short, these models allow for within class heterogeneity. Like the latent class model, this model posits distinguishable groups of people, but unlike the latent class model, also allows for within class heterogeneity, thereby encompassing both categorical and continuous aspects of latent structure. There are different variations of factor mixture models with different restrictions on the parameters. Here we fit a factor mixture model in which factor variances are fixed to zero within each class, factor means are free to vary across classes, and item thresholds and factor loadings are invariant across classes (this is also referred to as latent class factor analysis; Muthén, 2006).

All models were fit using Mplus, version 5 (Muthén & Muthén, 2007), using robust maximum likelihood estimation with a logit link function. An information-theoretic approach was used to determine the best fitting model from a series of alternatives (Campbell et al., 2009; Markon & Krueger, 2006). Specifically, we relied on the Bayesian information criterion (BIC; Schwartz, 1978) to compare model fit.[1]

Results

Prior to fitting any of the models, it was necessary to ensure the items were sufficiently unidimensional, meeting the models' assumption. Exploratory and confirmatory factor analyses were carried out. For both scales at both time points, the ratio of first to second eigenvalues exceeded 4.0, the TLI values exceeded .93, and the RMSEA values did not exceed .07. These each met the criteria indicating good fit.

When analyzed separately, the best fitting model was the same for males and females at both time points. Therefore, results are presented for males and females combined. Model comparisons for T_1 and T_2 can be found in Table 1. In Table 1 we note the number of free parameters for each model, the log-likelihood values, and BIC values. Lower BIC values indicate better fit to the data (for a more extensive discussion, see e.g., Krueger et al. 2005). We continued fitting latent class models with increasing numbers of classes until BIC began to increase, signaling a worse fit to the data. At both time points, the best fitting latent class model for aggression was a five class model. For rule breaking, the best fitting latent class model included three distinct classes. To determine the number of factor mixture models to fit, we first determined the best fitting latent class model, then fit up to that many classes in the factor mixture models (i.e., a factor mixture model with five classes for aggression and three classes for rule breaking). The best fitting factor mixture model for aggression at both time points was a five class model. For rule breaking, a two class model fit best at T_1, while a three class model fit best at T_2. Across the three types of models (latent trait, latent class, and factor mixture models), for both scales at both time points, the best fitting model was the latent trait model.

We inspected the item parameter estimates from the best fitting model (the latent trait model parameterization in a 2-PL metric). Item parameters define item information functions, which can be added to form an information function indicating the amount and location of information an entire test provides. Test information functions for aggression at T_1 and T_2 can be found in Figure 1. The latent trait (labeled “theta”) lies along the x axis. The amount of information the test provides, which is reciprocal to the standard error of the scale, can be traced along the y axis. Higher discrimination parameters correspond to more information (i.e., more area under the curve). The location parameters determine where along the latent trait the information is provided. With regards to aggression, at T_1 most items had high discrimination parameters, indicating they are good markers of aggression and can discriminate among individuals at varying levels of theta. The discrimination parameters ranged from .50 to 1.08 (M = .83). The location parameters fell in the moderate range, meaning the scale provides little information at either end of the aggression spectrum and is a stronger measure for individuals at moderate levels of aggression. The items with the lowest location parameters involved arguing a lot (b = -1.57) and being stubborn and irritable (b = -1.19). Those with the highest location parameters entailed destroying others' property (b = 1.52) and threatening people (b = 2.03). The parameters at T_2 were largely similar, as depicted in Figure 1. A complete list of the item parameters can be obtained from the first author.

In terms of rule breaking (see Figure 2), most items had high discrimination parameters. These ranged from .35 to 1.46 (M = .81) at T_1. Although fairly good at discrimination, the rule breaking items were limited in terms of their coverage of the latent trait continuum. Most location parameters were quite high, meaning the items are better markers of highly delinquent behavior as opposed to less delinquent behavior. The items with the lowest location parameters entailed feeling guilty after misbehaving (b = -.70) and using obscene language (b = -.69). The items with the highest location parameters revered to being truant (b = 3.14) and illegal drug use (b = 3.18). This pattern was replicated at T_2.

Finally, we investigated the stability of the latent factors over time. We fit a confirmatory latent trait model with two latent traits, which were T_1 and T_2 aggression. The correlation between the two was .72. The same model was fit for rule breaking, and the correlation between T_1 and T_2 was also .72.

Discussion

Our primary objective was to compare the relative fit of latent trait, latent class, and factor mixture models of externalizing behaviors in a large, representative sample of adolescents. Aggressive and delinquent behaviors were assessed with parent- and self-reports on the Child Behavior Checklist (Achenbach, 1991a; 1991b) at age 11 and again roughly 2.5 years later. At both time points, the best-fitting models were latent trait models, indicating that externalizing behaviors in adolescence are better conceptualized in dimensional terms.

Though not a direct replication, findings from our study support those from prior research. Krueger et al. (2005) and Markon and Krueger (2005) also found superior fit for a latent trait model in a study of externalizing diagnoses in adulthood. In the adolescent literature, latent trait modeling of externalizing behaviors has revealed meaningful latent traits (Gelhorn et al., 2009; Schrum & Salekin, 2006; Van den Oord et al., 2003), which supports a dimensional conceptualization. It should be noted that support for dimensional structures has even been found when modeling categorical symptoms and diagnoses (e.g., Gelhorn et al., 2009; Krueger et al., 2005; Markon & Krueger, 2005).

Although latent class analyses of adolescent externalizing behaviors involve a different model (Brownfield & Sorenson, 1987; Eaves et al., 1993; Nock et al., 2006; Odgers et al., 2007), findings from latent class research can be reconciled with ours. In each of the latent class analyses reported to date, the classes that emerged were graded in terms of severity, as opposed to being distinct in a nominal sense (i.e., classes where probabilities of item endorsement did not increase monotonically across classes but were instead unique in each class). Classes with distinct sets of symptoms or behavioral patterns would lend more support to a categorical perspective, though this is not what has been observed with latent class modeling of externalizing behaviors. For example, Brownfield and Sorenson’s (1987) findings suggested three groups including non-delinquent, moderately delinquent, and severely delinquent adolescents. Nock et al. (2006) concluded that the best fitting model included six ranked latent classes ranging from no conduct disorder to pervasive conduct disorder. A similar pattern was reported by Eaves et al. (1993) and Odgers et al. (2007). In short, each of these studies involved fitting latent class models to externalizing behaviors, but the classes that emerged seemed to represent degrees of severity on an underlying continuum. The latent trait model formally tests the ordering of these classes by making the latent variable continuous. While prior research involving both latent trait and latent class modeling has suggested that externalizing behaviors in adolescence are dimensional, our findings now provide firmer evidence.

The current findings should be interpreted within the context of the limitations of the study. First, we have focused on externalizing as defined by Achenbach (1991a), with emphasis on delinquent and aggressive behaviors. Some findings suggest that two latent factors account for externalizing behaviors; the first reflects oppositional behavior disorders, including attention deficit hyperactivity disorder and oppositional defiant disorder, while the second reflects social norm violation disorders, including conduct disorder, adult antisocial disorder, and substance use disorders (Farmer, Seeley, Kosty, & Lewinsohn, 2009). In this study, we focus exclusively on syndromes related to this second factor. The latent structure of syndromes related to the first factor has been investigated in prior work (Lubke et al., 2007). Second, the conclusions only reflect early adolescence, from roughly 11 to 13.5 years of age. Third, we have only two time points available, limiting our ability to model latent growth over time. We have demonstrated that within these two time points, externalizing behaviors are distributed continuously. Nevertheless, cross sectional continuity of externalizing problems does not preclude distinct trajectories over time. For example, distinct groups could emerge based on age of onset or duration of involvement, as Moffitt (1993) discussed. Indeed, in several latent class growth curve analyses of adolescent externalizing behaviors, distinct classes based on different growth trajectories have emerged (e.g., Lee & Thompson, 2009; Martino, Ellickson, Klein, McCaffrey, & Edelen, 2008; Reinecke, 2006). The best classification across time is conceptually separate from the best classification within particular time points. With only two time points, however, we were unable evaluate across time classification, as a minimum of three time points are needed to identify latent growth curve models (Duncan & Duncan, 2004). With additional time points, we could investigate whether differences in type, onset, and/or trajectory of externalizing behaviors can be best modeled with distinct groups (e.g., Moffitt, 1993), or whether differences can be best modeled with continuous distributions (Lahey & Waldman, 2003). A final point to consider for future research involves model estimation and evaluation of model fit. Alternative and newer approaches to model estimation and evaluation of model fit represent important potential extensions of the approaches we took here. For example, alternative approaches to evaluating model fit, such as posterior predictive checks (e.g., Loken, 2004), are possible using Bayesian analysis and Markov Chain Monte Carlo estimation methods. This is an important direction for future research on comparing categorical and continuous representations of latent structure.

Despite these limitations, this study is the first to directly compare the relative fit of latent trait, latent class, and factor mixture models of externalizing behaviors at two time points in adolescence, and we provide evidence that externalizing behaviors in adolescence are dimensional (Hudziak, Achenbach, Althoff, & Pine, 2009). This is important because it shows how classification models can be based on data rather than a priori preferences. Currently, most major psychiatric classification systems conceptualize externalizing problems in terms of discrete groups. Although thresholds can always be set on continua to identify groups, such groups may be better conceptualized as differing along a continuum (Lahey et al., 1994; Lahey & Waldman, 2003). The transition from *DSM-IV* to *DSM-5* provides an important opportunity to better incorporate the evidence for the dimensionality of psychopathology into official nosologies. Indeed, this has been described by the DSM-5 task force chair and co-chair as a major aim of the revision effort (Regier, Narrow, Kuhl, & Kupfer, 2009).

Furthermore, the importance of assessing symptoms along entire externalizing continua is emphasized by our findings. Indicators of externalizing behaviors often target a narrow portion of a continuum (e.g., Cooke & Michie, 1997; Gelhorn et al., 2009; Van den Oord, 2003; Walton, Roberts, Krueger, Blonigen, & Hicks, 2008). For example, as shown here, indicators of rule breaking behavior are most sensitive to individuals with a high tendency for those behaviors. Recognizing the dimensional nature of these behaviors, we can aim to augment conceptualization and assessment to include less severe forms of externalizing, thereby providing a more comprehensive account of these costly forms of behavior problems.

References

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 profile.* Burlington: University of Vermont.

Achenbach, T. M. (1991b). *Manual for the Youth Self-Report and 1991 profile.* Burlington: University of Vermont.

Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin*, *131*, 361-382.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Brownfield, D., & Sorsenson, A. M. (1987). A latent structure analysis of delinquency. *Journal of Quantitative Criminology, 3*, 103-124.

Campbell, S. B., Morgan-Lopez, A. A., Cox, M. J., McLoyd, V. C., & National Institute of Child Health and Human Development Early Child Care Research Network. (2009). A latent class analysis of maternal depressive symptoms over 12 years and offspring adjustment in adolescence. *Journal of Abnormal Psychology, 118*, 479-493.

Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion. (2007). *Healthy Youth! YRBSS. National Trends in Risk Behavior*. Retrieved from http://www.cdc.gov/HealthyYouth/yrbs/trends.htm

Cooke, D. J., & Michie, C. (1997). An item response theory analysis of the Hare Psychopathy Checklist-Revised. *Psychological Assessment, 9*, 3-14.

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.

De Winter, A. F., Oldehinkel, A. J., Veenstra, R., Brunnekreef, J. A., Verhulst, F. C., & Ormel, J. (2005). Evaluation of non-response bias in mental health determinants and outcomes in a large sample of preadolescents. *European Journal of Epidemiology*, *20*, 173-181.

Dick, D. M., Aliev, F., Wang, J. C., Grucza, R. A., Schuckit, M., Kuperman, S., … Goate, A. (2008). Using dimensional models of externalizing psychopathology to aid in gene identification. *Archives of General Psychiatry*, *65*, 310-318.

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, *35*, 333-363.

Eaves, L. J., Silerg, J. L., Hewitt, J. K., Rutter, M., Meyer, J. M., Neale, M. C., & Pickles, A. (1993). Analyzing twin resemblance in multisymptom data: Genetic applications of a latent class model for symptoms of conduct disorder in juvenile boys. *Behavior Genetics*, *23*, 5-19.

Farmer, R. F., Seeley, J. R., Kosty, D. B., & Lewinsohn, P. M. (2009). Refinements in the hierarchical structure of externalizing psychiatric disorders: Patterns of lifetime liability from mid-adolescence through early adulthood. *Journal of Abnormal Psychology*, *118*, 699-710.

Frick, P. J., Cornell, A. H., Barry, C. T., Bodin, S. D., & Dane, H. E. (2003). Callous-unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report of delinquency. *Journal of Abnormal Child Psychology*, *31*, 457-470.

Gelhorn, H., Hartman, C., Sakai, J., Mikulich-Gilbertson, S., Stallings, M., Young, S., … Crowley, T. (2009). An item response theory analysis of DSM-IV conduct disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*, 42-50.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Publications, Inc.

Hudziak, J. J., Achenbach, T. M., Althoff, R. R., & Pine, D. S. (2009). A dimensional approach to developmental psychopathology. In J. E. Helzer, H. C. Kraemer, R. F. Krueger, H. Wittchen, P. S. Sirovatka, & D. A. Regier (Eds.), *Dimensional approaches in diagnostic classification: Refining the research agenda for DSM-V* (pp. 101-113). Arlington, VA: American Psychiatric Association.

Huisman, M, Oldehinkel, A. J., de Winter, A., Minderaa, R. B., de Bildt, A., Huizink, A. C., Verhulst, F. C., & Ormel, J. (2008). Cohort Profile: The Dutch 'Tracking Adolescents' Individual Lives Survey'; TRAILS. *International Journal of Epidemiology*, *37*, 1227-1235.

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, *60*, 929-937.

Krueger, R. F., Markon, K. E., Patrick, C. J., & Iacono, W. G. (2005). Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implications for DSM-V. *Journal of Abnormal Psychology*, *114*, 537-550.

Lahey, B. B., et al. (1994). DSM-IV field trials for oppositional defiant disorder and conduct disorder in children and adolescents. *The American Journal of Psychiatry*, *151*, 1163-1171.

Lahey, B. B., & Waldman, I. D. (2003). A developmental propensity model of the origins of conduct problems during childhood and adolescence. In B. B. Lahey, T. E. Moffitt, & A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 76-117). New York: The Guilford Press.

Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. *Developmental Psychology*, *44*, 446-456.

Laub, J., & Sampson, R. (1994). Unemployment, marital discord, and deviant behavior: The long-term correlates of childhood misbehavior. In T. Hirschi & M. Gottfredson (Eds.), *The generality of deviance* (pp. 133-161). New Brunswick, NJ: Transaction Publishers.

Lee, B. R., & Thompson, R. (2009). Examining externalizing behavior trajectories of youth in group homes: Is there evidence for peer contagion? *Journal of Abnormal Child* Psychology, 37, 31-44.

Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1991). Differences and similarities between children, mothers, and teachers as informants on disruptive child behavior. *Journal of Abnormal Child Psychology*, *19*, 75-95.

Loken, E. (2004). Using latent class analysis to model temperament types. *Multivariate Behavioral Research*, *39*, 625-652.

Lubke, G. H., Muthén, B., Moilanen, I. K., McGough, J. J., Loo, S. K., Swamson, J. M., … Smalley, S. S. (2007). Subtypes versus severity differences in attention-deficit/hyperactivity disorder in the Northern Finnish Birth Cohort. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 1584-1593.

Lubke, G., & Neale, M. C. (2006). Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood. *Multivariate Behavioral Research*, *41*, 499-532.

Markon, K. E., & Krueger, R. F. (2005). Categorical and continuous models of liability to externalizing disorders. *Archives of General Psychiatry*, *62*, 1352-1359.

Markon, K. E., & Krueger, R. F. (2006). Information-theoretic latent distribution modeling: Distinguishing discrete and continuous latent variable models. *Psychological Methods*, *3*, 228-243.

Martino, S. C., Ellickson, P. L., Klein, D. J., McCaffrey, D., & Edelen, M. O. (2008). Multiple trajectories of physical aggression among adolescent boys and girls. *Aggressive Behavior*, *34*, 61-75.

McCutcheon, A. L. (1987). *Latent Class Analysis*. Newbury Park, CA: Sage Publications, Inc.

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

Monshouwer, K., Verdurmen, J., van Dorsselaer S., Smit E., Gorter, A., & Vollebergh, W. (2008). *Jeugd en riskant gedrag 2007. Kerngegevens uit het peilstationsonderzoek scholieren* [*Adolescents and risk-taking behavior 2007*]. Utrecht: Trimbos Instituut.

Muthén, B. (2006). Should substance use disorders be considered as categorical or dimensional? *Addiction*, *101*, 6-16.

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

Nock, M. K., Kazdin, A. E., Hiripi, E., & Kessler, R. C. (2006). Prevalence, subtypes, and correlates of DSM-IV conduct disorder in the National Comorbidity Survey Replication. *Psychological Medicine*, *36*, 699-710.

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.

Odgers, C. L., Moretti, M. M., Burnette, M. L., Chauhan, P., Waite, D., & Reppucci, N. D. (2007). A latent variable modeling approach to identifying subtypes of serious and violent female juvenile offenders. *Aggressive Behavior*, *33*, 339-352.

Piacentini, J. C., Cohen, P., & Cohen, C. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology*, *20*, 51-63.

Regier, D. A., Narrow, W. E., Kuhl, E. A., & Kupfer, D. J. (2009). The conceptual development of DSM-V. *American Journal of Psychiatry, 166,* 645-650.

Reinecke, J. (2006). Longitudinal analysis of adolescents' deviant and delinquent behavior: Applications of latent class growth curves and growth mixture models. *Methodology*, *2,* 100-112.

Schrum, C. L., & Salekin, R. T. (2006). Psychopathy in adolescent female offenders: an item response theory analysis of the psychopathy checklist: Youth version. *Behavioral Sciences and the Law*, *24*, 39-63.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461-464.

Siennick, S. E. (2007). The timing and mechanisms of the offending-depression link. C*riminology: An Interdisciplinary Journal, 45,* 583-615.

Tanner, J., Davies, S., & O'Grady, B. (1999). Whatever happened to yesterday's rebels? Longitudinal effects of youth delinquency and education and employment. *Social Problems*, *46*, 250-274.

U.S. Department of Justice, Federal Bureau of Investigation. (2009). *Crime in the United States, 2008*. Retrieved from http://www.fbi.gov/ucr/cius2008/arrests/index.html

Van den Oord, E. J. C. G., Pickles, A., & Waldman, I. D. (2003). Normal variation and abnormality: an empirical study of the liability distributions underlying depression and delinquency. *Journal of Child Psychology and Psychiatry*, *44*, 180-192.

Verhulst, F. C., Van der Ende, J., & Koot, H. M. (1996). *Handleiding voor de CBCL/4-18*. Afdeling Kinder en Jeugdpsychiatrie, Sophia Kinderziekenhuis/Academisch Ziekenhuis Rotterdam/Erasmus Universiteit Rotterdam.

Walton, K. E., Roberts, B. W., Krueger, R. F., Blonigen, D. M., & Hicks, B. M. (2008). Capturing abnormal personality with normal personality inventories: An item response theory approach. *Journal of Personality*, *76*, 1623-1648.

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition. *Journal of Abnormal Psychology*, *114*, 494-504.

Footnote

[1] Mplus computes three information theoretic indices, but two of these (AIC and sample size adjusted BIC) have been shown to lead to incorrect conclusions in identifying the correct latent variable model. In contrast, BIC was shown to perform better in simulation studies in terms of correctly identifying the true population model (Nylund, Asparouhov, & Muthén, 2007).

Table 1

Summary of the Fit of Latent Trait, Latent Class, and Factor Mixture Models for Aggression and Rule Breaking at Time 1 and Time 2

Model	*k*	BIC	ln(*L*)
T_1 Aggression			
Latent trait	**34**	**35443.925**	**-17592.519**
Latent class			
1 class	17	40272.817	-20071.687
2 classes	35	36393.101	-18063.300
3 classes	53	35692.312	-17644.376
4 classes	71	35570.848	-17515.116
5 classes	89	35529.371	-17425.849
6 classes	107	35550.797	-17368.033
Factor mixture			
2 classes	36	36400.716	-18063.301
3 classes	37	35624.234	-17671.252
4 classes	39	35504.963	-17604.002
5 classes	41	35499.979	-17593.896
T_1 Rule Breaking			
Latent trait	**30**	**24200.852**	**-11986.212**
Latent class			
1 class	15	25934.337	-12952.169
2 classes	31	24486.677	-12125.317
3 classes	47	24314.918	-11978.523
4 classes	63	24375.391	-11947.845
Factor mixture			
2 classes	32	24494.292	-12125.317
3 classes	33	24241.188	-11994.958
T_2 Aggression			
Latent trait	**34**	**31338.790**	**-15541.142**
Latent class			
1 class	17	35966.062	-17918.904
2 classes	35	32175.914	-15955.931
3 classes	53	31632.277	-15616.214
4 classes	71	31554.074	-15509.213
5 classes	89	31523.334	-15425.944
6 classes	107	31541.094	-15366.925
Factor mixture			
2 classes	36	32183.460	-15955.932
3 classes	37	31559.050	-15639.955
4 classes	39	31413.187	-15559.479
5 classes	41	31402.033	-15546.358

T_2 Rule Breaking			
Latent trait	**30**	**23294.113**	**-11533.892**
Latent class			
1 class	15	25557.847	-12722.341
2 classes	31	23653.893	-11710.009
3 classes	47	23440.709	-11543.063
4 classes	63	23447.61	-11486.159
Factor mixture			
2 classes	32	23661.442	-11710.012
3 classes	33	23346.541	-11548.789

Note. k = number of parameters. BIC = Bayesian information criterion. ln (L) = log-likelihood. T_1 = Time 1. T_2 = Time 2. The best fitting model for each time point and trait is presented in bold.

Figure 1

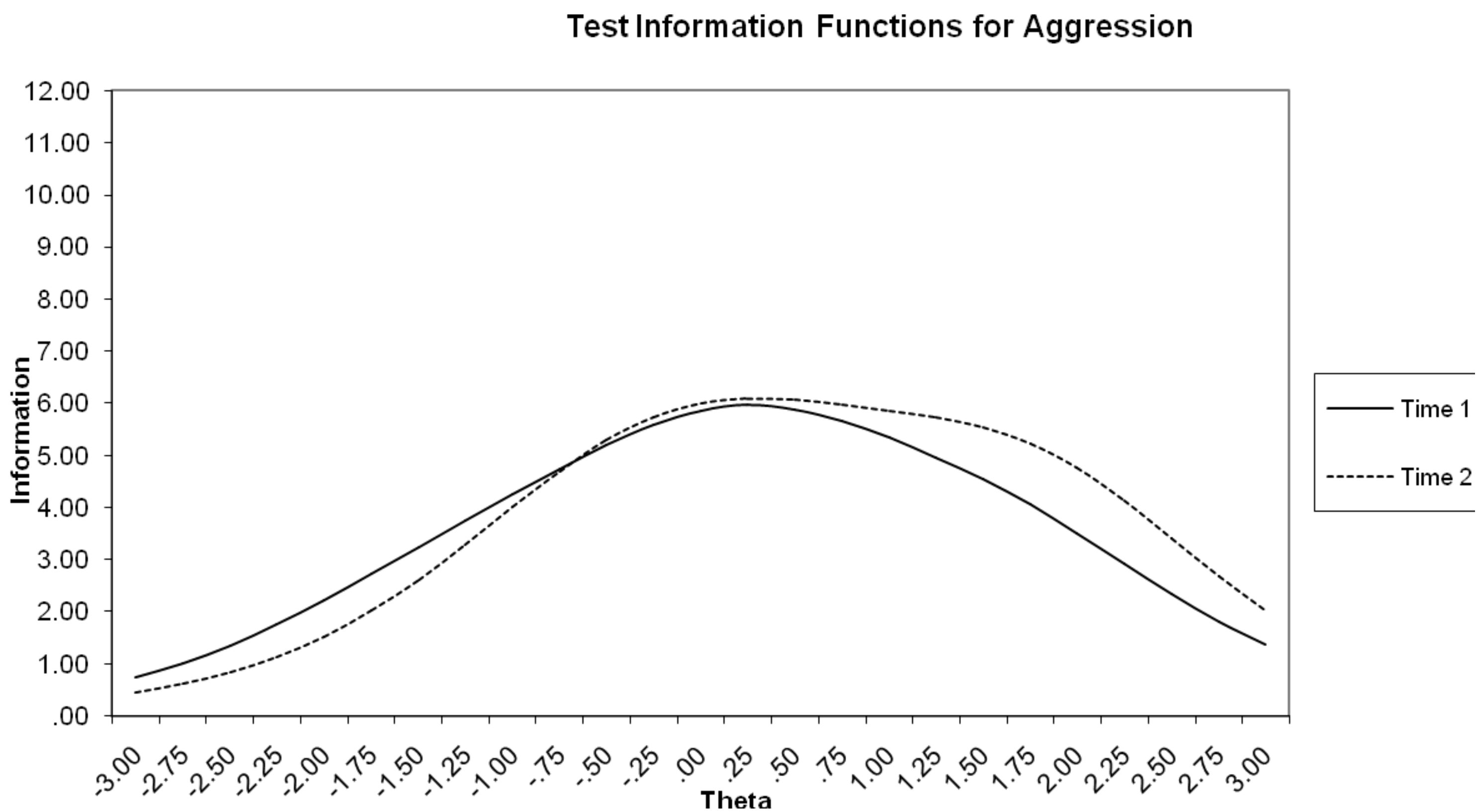
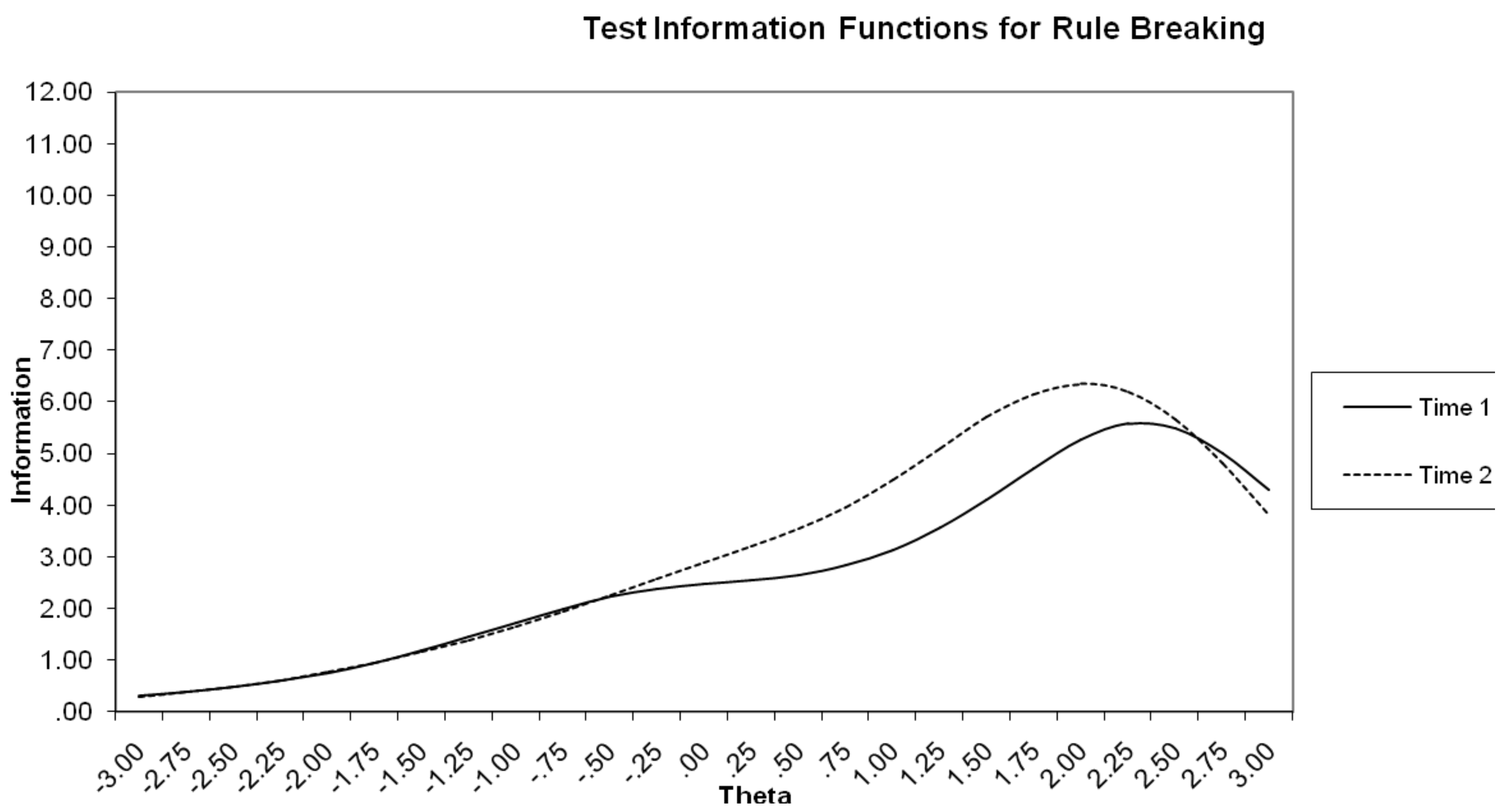

Test Information Functions for Aggression
Information
12.00
11.00
10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
.00
-3.00
-2.75
-2.50
-2.25
-2.00
-1.75
-1.50
-1.25
-1.00
-.75
-.50
-.25
.00
.25
.50
.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Theta
Time 1
Time 2

Figure 2


Test Information Functions for Rule Breaking
12.00
11.00
10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
.00
Information
-3.00
-2.75
-2.50
-2.25
-2.00
-1.75
-1.50
-1.25
-1.00
-.75
-.50
-.25
.00
.25
.50
.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Theta
Time 1
Time 2