



Published in final edited form as:

J Abnorm Child Psychol. 2013 February ; 41(2): 223–237. doi:10.1007/s10802-012-9674-z.

Evidence for a Multi-Dimensional Latent Structural Model of Externalizing Disorders

Katie Witkiewitz,

Department of Psychology, Washington State University, 14204 NE Salmon Creek Ave, Vancouver, WA 98686, USA, katie.witkiewitz@wsu.edu

Kevin King,

Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195-1525, USA, kingkm@uw.edu

Robert J. McMahon,

Department of Psychology, Simon Fraser University and the Child & Family Research Institute, 8888 University Drive, Burnaby, BC V5A 1S6, Canada, rjcmcmah@sfu.ca

Johnny Wu,

Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195-1525, USA, johnnywu@uw.edu

Jeremy Luk, and

Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195-1525, USA, jwluk@uw.edu

the Conduct Problems Prevention Research Group

Abstract

Strong associations between conduct disorder (CD), antisocial personality disorder (ASPD) and substance use disorders (SUD) seem to reflect a general vulnerability to externalizing behaviors. Recent studies have characterized this vulnerability on a continuous scale, rather than as distinct categories, suggesting that the revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) take into account the underlying continuum of externalizing behaviors.

© Springer Science+Business Media, LLC 2012

Correspondence to: Katie Witkiewitz.

Members of the Conduct Problems Prevention Research Group include Karen L. Bierman (Pennsylvania State University), Kenneth A. Dodge (Duke University), Mark T. Greenberg (Pennsylvania State University), John E. Lochman (University of Alabama), Robert J. McMahon (Simon Fraser University and Child & Family Research Institute), and Ellen E. Pinderhughes (Tufts University). Katie Witkiewitz is now at the University of New Mexico. 1 University of New Mexico MSC03-2220, Albuquerque, NM 87131, (505) 277-4121, katiew@unm.edu

Members of the Conduct Problems Prevention Research Group, in alphabetical order, include Karen L. Bierman, Department of Psychology, Pennsylvania State University; John D. Coie, Duke University; Kenneth A. Dodge, Center for Child and Family Policy, Duke University; Mark T. Greenberg, Department of Human Development and Family Studies, Pennsylvania State University; John E. Lochman, Department of Psychology, the University of Alabama; Robert J. McMahon, Department of Psychology, Simon Fraser University, and the Child & Family Research Institute; and Ellen E. Pinderhughes, Department of Child Development, Tufts University.

Drs. Bierman, Coie, Dodge, Greenberg, Lochman, and McMahon are the developers of the Fast Track curriculum and have a publishing agreement with Oxford University Press. Dr. Greenberg is an author on the PATHS curriculum and has a royalty agreement with Channing-Bete, Inc. Dr. Greenberg is a principal in PATHS Training, LLC. Dr. McMahon is a coauthor of *Helping the Noncompliant Child* and has a royalty agreement with Guilford Publications, Inc.; he is also a member of the Treatments That Work Scientific Advisory Board with Oxford University Press.

The authors are grateful for the collaboration of the school districts that participated in the Fast Track Project as well as the hard work and dedication of the many staff members who implemented the project, collected the evaluation data, and assisted with data management and analyses. For additional information concerning Fast Track, see <http://www.fasttrackproject.org>.

However, most of this research has not included measures of disorders that appear in childhood [e.g., attention-deficit/hyperactivity disorder (ADHD) or oppositional defiant disorder (ODD)], nor has it considered the full range of possibilities for the latent structure of externalizing behaviors, particularly factor mixture models, which allow for a latent factor to have both continuous and categorical dimensions. Finally, the majority of prior studies have not tested multidimensional models. Using lifetime diagnoses of externalizing disorders from participants in the Fast Track Project ($n = 715$), we analyzed a series of latent variable models ranging from fully continuous factor models to fully categorical mixture models. Continuous models provided the best fit to the observed data and also suggested that a two-factor model of externalizing behavior, defined as (1) ODD+ADHD+CD and (2) SUD with adult antisocial behavior sharing common variance with both factors, was necessary to explain the covariation in externalizing disorders. The two-factor model of externalizing behavior was then replicated using a nationally representative sample drawn from the National Comorbidity Survey-Replication data ($n = 5,692$). These results have important implications for the conceptualization of externalizing disorders in DSM-5.

Keywords

Externalizing; Classification; Mixture modeling; Substance abuse; Substance dependence

The current versions of the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition–Text Revision (DSM-IV-TR; American Psychiatric Association 2000) and the International Classification of Diseases (ICD-10; World Health Organization 1992) require that each type of problem behavior be considered as a distinct psychiatric disorder (e.g., attention-deficit/hyperactivity disorder [ADHD], oppositional defiant disorder [ODD]). The DSM and ICD categorical diagnostic systems, which were developed in the latter part of the 20th century, greatly advanced the field by providing objective criteria for characterizing psychiatric illnesses (First 2010).

Yet, recent research has suggested that most forms of psychopathology occur more dimensionally than categorically (Widiger and Samuel 2005). Moreover, current classification systems do not provide a system for diagnosing individuals who have symptoms across multiple diagnostic categories, and critics of the current classification systems have suggested dimensional models of psychopathology may provide a better understanding of psychiatric disorders (e.g., Kamphuis and Noordhof 2009; Krueger and Bezdjian 2009). Recently, Markon and colleagues (2011) reviewed 58 studies with over 59,000 participants and concluded that continuous (i.e., dimensional) measures of psychopathology provide a 15 % increase in reliability and a 37 % increase in validity, in comparison to categorical measures. However, as noted by Brown and Barlow (2005), there remain numerous questions as to how a dimensional system might be developed and how such a system could be introduced in DSM-5 (American Psychiatric Association 2010). Brown and Barlow (2005) further argued that latent variable modeling may be one way forward.

Latent variable models encompass a wide variety of statistical models that can be used to model observed measures (i.e., manifest variables) as indicators of an underlying, unobservable construct (i.e., latent variables). Psychiatric diagnoses are analogous to latent variable models in that observed behavioral symptoms are used as indicators of the underlying, unobservable diagnosis. Both latent variables and diagnoses are defined by the observable measures that are considered relevant to the underlying latent measure or diagnostic category, where the inclusion of particular observable measures or symptoms defines the structure of the latent variable or diagnosis.

Importantly, the distributions of latent variables can be categorical (often referred to as latent classes or finite mixtures), dimensional (i.e., continuous, referred to as factors or traits), or a hybrid of categorical and dimensional (i.e., factor mixture models; Masyn et al. 2010). Hybrid models test whether there is dimensional severity within distinct categories, whether distinct categories are quantifiable by dimensional severity, or any other combination of dimensional and categorical measurement of the underlying construct. In many ways, the hybrid models are most consistent with a diagnostic system that is based on dimensional symptom profiles. The current DSM is analogous to a hybrid model in which the latent categorical diagnosis is derived from a number of symptoms, often along multiple dimensions.

Dimensional Models of Externalizing Disorders

One of the dimensional assessment proposals, put forward by Krueger and colleagues (Krueger and Bezdjian 2009; Krueger et al. 2005), was to create a hierarchical system for characterizing externalizing disorders (e.g., CD, antisocial personality disorder (ASPD), and substance use disorders (SUD)) that occur along a continuum of risk for externalizing symptomatology (e.g., physical aggression, problematic substance use). Specifically, Krueger and South (2009) proposed that CD, substance dependence, and ASPD be considered together as an externalizing disorder classification cluster. Their proposal is largely based on the high degree of comorbidity among externalizing behaviors that can be explained by a latent, genetically mediated vulnerability for externalizing disorders (see Krueger et al. 2002, 2005, 2007). Several studies using latent variable models have provided empirical evidence that shared variance between these disorders can be characterized as a single, unidimensional continuous latent variable (Krueger et al. 2007; Markon and Krueger 2005; Tuvblad et al. 2009). Yet, some have questioned whether it is appropriate to consider CD, ASPD, and SUD as part of an externalizing disorders classification cluster (Jablensky 2009) or whether alternative structures should be considered. It could be the case that a multidimensional factor model or a hybrid model is necessary to explain the covariation in externalizing disorders, and prior studies of externalizing disorders have not tested these alternative models (Krueger et al. 2005; Markon and Krueger 2005; Tuvblad et al. 2009).

More importantly, the externalizing spectrum models examined in the literature have largely focused on adult psychopathology and have not included the childhood diagnoses or symptoms of ODD and ADHD, which have long been described as externalizing behaviors (Achenbach 1966). The omission of these childhood diagnoses from the externalizing spectrum is problematic because these models have ignored the fact that adult externalizing disorders (e.g., SUD, ASPD) can often be traced back to externalizing behavior problems in early childhood, oftentimes via a developmental sequence (e.g., Beauchaine et al. 2010; Loeber and Burke 2011). Given the strong association between child and adult externalizing disorders, it could be the case that measurement models of the externalizing spectrum that have not included childhood diagnoses or symptoms of ODD and ADHD (e.g., Markon and Krueger 2005) may have been misspecified. An important extension of this prior work is to test the externalizing spectrum with the inclusion of childhood diagnoses.

Recent work by Farmer and colleagues (2009) that did include childhood diagnoses of externalizing disorders concluded that a two-factor model (Fig. 1, model 2b) that distinguished “oppositional behaviors” from “social norm violation behaviors” provided the best fit to the data. Alternative two- and three-factor models (Fig. 1, models 2a and 3) also provided a better fit to the data than a one-factor model. Similarly, Verona et al. (2011) provided support for a two-dimensional model of adolescent psychopathology in a sample of 10–17 year olds that included externalizing (CD, ADHD, and ODD) and substance use (alcohol and marijuana) factors. Based on all of this research, the one-factor model of

externalizing disorders, defined by CD, ASPD, and SUD, advocated by Krueger and colleagues (Krueger et al. 2005; Markon and Krueger 2005; Walton et al. 2011) may not be sufficient to explain the covariation in externalizing disorders, particularly when ADHD and ODD are included in the model.

Excluding or including a particular disorder in a model of externalizing disorders is particularly important given the nature of latent variable models, where the inclusion or exclusion of observed indicators can change the structure, dimensionality, and interpretation of the latent variable. Latent variables represent the shared variance among observed measures; thus, the meaning of the latent variable changes depending on which observed measures are utilized. Model fit indices tell us how well the specified latent factor model reproduces the observed covariation among measures. Under ideal circumstances the model provides both excellent fit to the data, and the latent variable explains a significant amount of the variance in the observed measures. Latent variables almost never explain 100% of the variance in the observed measures, and the residual, unexplained variance in the observed measures is assumed to be comprised of both random error and systematic variability in the measure that is not explained by the latent factor. For example, Lahey and colleagues (2008) found that 68% to 82 % of the variance in ADHD, CD, and ODD factors was explained by a higher-order externalizing factor, suggesting a high degree of covariation that explained a large proportion of variance in each of the individual disorders. At the same time, 18 % to 32 % of the variance in the individual disorders was not explained by the externalizing factor, suggesting differentiation among disorders (Lahey et al. 2011).

Current Study

The main goal of the current study was to explore multiple models of this shared variation among externalizing disorders. We attempted to replicate and extend prior research by Krueger and colleagues (Krueger et al. 2005; Markon and Krueger 2005) by assessing the covariation between conduct problems, substance use, and adult antisocial behavior, as well as diagnoses of ADHD and ODD, along the dimensional-categorical spectrum using data from the Fast Track project (Conduct Problems Prevention Research Group 1992, 2000). We approached the question of whether externalizing behavior is continuous or categorical using a series of latent structure models along the dimensional-categorical spectrum described by Masyn and colleagues (2010). Importantly, the current study extends the work of Krueger and colleagues (e.g., Krueger et al. 2005; Markon and Krueger 2005) by adding measures of externalizing behaviors that are more common during childhood (ODD and ADHD) and by testing additional models along the dimensional-categorical spectrum. Given that latent structure models can be susceptible to sample-specific characteristics, we conducted a replication of the best-fitting models in a separate nationally representative sample from the National Comorbidity Survey-Replication (NCS-R; Kessler et al. 2004).

Methods

Participants

Fast Track—Participants came from the control schools of a longitudinal multi-site investigation of the development and prevention of childhood conduct problems, the Fast Track project (Conduct Problems Prevention Research Group 1992, 2000). Schools within four sites (Durham, NC, Nashville, TN, Seattle, WA, and rural Pennsylvania) were identified as high risk based on crime and poverty statistics of the neighborhoods that they served. Within each site, schools were divided into sets matched for demographics (size, percentage free or reduced lunch, ethnic composition), and the sets were randomly assigned to control and intervention groups. Using a multiple-gating screening procedure that combined teacher and parent ratings of disruptive behavior, 9,594 kindergarteners across

three cohorts (1991–93) from 55 schools were screened initially for classroom conduct problems by teachers, using the Teacher Observation of Child Adjustment-Revised (TOCA-R) Authority Acceptance score (Werthamer-Larsson et al. 1991). Those children scoring in the top 40 % within cohort and site were then solicited for the next stage of screening for home behavior problems by the parents, using items from the Child Behavior Checklist (Achenbach 1991) and similar scales, and 91 % agreed ($n = 3,274$). The teacher and parent screening scores were then standardized and summed to yield a total severity-of-risk screen score. Children were selected for inclusion into the high-risk sample based on this screen score, moving from the highest score downward until desired sample sizes were reached within sites, cohorts, and groups. Deviations were made when a child failed to matriculate in the first grade at a core school ($n = 59$) or refused to participate ($n = 75$), or to accommodate a rule that no child would be the only girl in an intervention group. The outcome was that 891 children (control = 446, intervention = 445) participated. In addition to the high-risk sample of 891, a stratified normative sample of 387 children was identified to represent the population normative range of risk scores and was followed over time. From among the control schools ($n = 27$), teachers completed ratings of child disruptive behavior to identify a normative, within-site stratified sample of about 10 children within each decile of behavior problems. Follow-up assessments were conducted annually through 2 years post high school (approximately age 20).

Across time, an average of 90 % of participants were retained at each time point, and prior analyses of these data suggested that participants lost to follow-up did not significantly differ from those retained (Conduct Problems Prevention Research Group 1999). Of particular concern for the current study was missing data in the later years of assessment, particularly the assessments of substance use and antisocial behavior. Substance use diagnostic information in the 2 years post-high school was available for 602 participants (79.8 %) and assessment of antisocial behavior was available for 512 participants (67.9 %). Those with missing substance use data did not significantly differ from those retained on any demographic variables or any other externalizing diagnoses. Individuals with missing data on the measure of antisocial behavior did not significantly differ from those retained on age, sex, lifetime substance use diagnoses, or lifetime CD diagnoses. Those with missing antisocial data were significantly more likely to be African American ($\chi^2(2) = 13.2, p = 0.001$), have a lifetime ODD diagnosis ($\chi^2(1) = 5.43, p = 0.02$), and/or have a lifetime ADHD diagnosis ($\chi^2(1) = 4.62, p = 0.03$).

The current study utilized data from the high-risk control group (65 % male; 49 % African American, 48 % Euro American, 3 % other race) and normative sample (51 % male; 43 % African American, 52 % Euro American, 5 % other race). Because 79 of those recruited for the high-risk control group were also included as part of the normative sample, the sample for the current analyses were based on a total of 754 participants. Of the 754 participants, 39 individuals did not have any valid data for any of the lifetime diagnoses or diagnostic criterion counts and these 39 individuals were excluded from the models described below. Thus, the final sample size for the current analyses was 715 participants (mean age at start of study = 6.56 ($SD = 0.44$); 58.6 % male, 46.4 % African American, 50.2 % Euro American, and 3.4 % other race; majority lower to low middle class). Weighting was used in all analyses to reflect the over-sampling of high-risk children. Participants from the high-risk intervention sample were not included in this study.

National Comorbidity Survey-Replication (NCS-R)—In addition, we included data from the NCS-R. The NCS-R is a nationally representative household survey of the prevalence of DSM-IV mental disorders among English-speaking adults (18 and older) conducted between February 2001 and April 2003 (see Kessler et al. 2004, for more information). Data from the 5,692 individuals who also completed the drug module of the

Composite International Diagnostic Interview (CIDI; Kessler and Ustun 2004) were used in the current analyses. The sample was 53 % female, with ages ranging from 18 to 60 (average age = 44.99, $SD = 17.9$). The participants were 72.8% Caucasian, 12.4 % African-American, 11.1 % Hispanic, and 3.8 % from other ethnicities. Socioeconomic status was nationally representative with an average household income of \$35,732 ($SD = \$31, 236$). NCS-R statistical weights were used to ensure that the sample was representative of the United States population.

Diagnostic Interviews

Fast Track—The Parent and Child Interview versions of the NIMH Diagnostic Interview Schedule for Children (DISC) are well-validated, highly structured, computer-administered, clinical interviews to assess DSM symptoms of ODD, CD, and ADHD in children and adolescents ages 6 to 17 years. We used the Parent Interview version 2.3 in grade 3 and the Parent and Child Interview version IV in grades 6, 9, and 12 (Shaffer and Fisher 1997; Shaffer et al. 2003). Lay interviewers, blind to control/normative status, were trained in clinical methods and scoring accuracy until each interviewer had reached criteria for reliable scoring of the DISC. Administration took place in the child's home with the primary parent, usually the mother, during the summer following grades 3, 6, 9, and 12; interviews with the child took place during the summer following grades 6, 9, and 12. Variables were computed for lifetime diagnoses of CD, ODD, and ADHD across all years of administration with diagnoses for grade 3 based on DSM-III-R criteria, and diagnoses for grades 6, 9, and 12 based on DSM-IV criteria. Rates of lifetime diagnoses in the normative and high-risk control samples were: CD (18.6 % combined, 13.7 % of total normative sample), ODD (25.7 % combined, 18.1% of total normative sample), ADHD (21.8% combined, 12.1 % of total normative sample).

The DISC-Young Adult version (DISC-YA; Shaffer et al. 2000) was administered to youth at 1- and 2-years post-high school to assess substance abuse/dependence and at 2 years post-high school to assess current ASPD, CD, ODD, and ADHD. The CD, ODD, and ADHD diagnoses were based on the DSM-IV criteria. Adult antisocial behavior (AAB) was based on having three or more criteria of ASPD derived from seven antisocial symptom items, with dimensional scores ranging from 0 to 7 ($M = 0.93$, $SD = 1.46$). Importantly we did not use diagnosis of ASPD because it requires evidence of CD prior to age 15, making the diagnosis of ASPD confounded with the diagnosis of CD. Alcohol abuse or dependence (ALC), marijuana abuse or dependence (MJ), and other substance abuse or dependence (DRG; i.e., stimulants, opiates, sedatives, hallucinogens, inhalants, and other prescription drugs used non-medically) were calculated from 11 symptom items (4 abuse items and 7 dependence items). If participants met criteria for DSM-IV substance abuse or dependence then they were considered to have a diagnosis of a SUD. Rates of lifetime diagnoses in the normative and highrisk control samples were: AAB (11.9 % combined, 4.9 % of normative sample), ALC (15.0 % combined, 13.0 % of normative sample), MJ (14.5 %, 10.7 % of normative sample), DRG (2.8 %, 1.6 % of normative sample).

National Comorbidity Survey–Replication—The Composite International Diagnostic Interview (CIDI; Kessler and Ustun 2004) was used to derive lifetime diagnoses of CD, ODD, ADHD, alcohol abuse/dependence (ALC), and drug abuse/dependence including marijuana abuse/dependence (DRGMJ). Marijuana abuse/dependence was included in the DRGMJ diagnosis because in the NCS-R data it is not possible to determine abuse/dependence criteria for each individual drug type. The CIDI is a comprehensive, fully-structured interview that was administered by trained interviewers for the assessment of mental disorders according to the definitions and criteria of ICD-10 and DSM-IV. All interviews were conducted face-to-face in the participant's home. For the purposes of the

current study, a lifetime adult antisocial behavior disorder variable was created from eight personality items in the NCS-R interview that represented five (deceitfulness, impulsivity, aggressiveness, reckless disregard, lack of remorse) of the criteria for ASPD. Individuals who endorsed at least four of the five ASPD criteria in the NCS-R were designated as engaging in adult antisocial behavior (AAB) for the current study. Rates of lifetime diagnoses and adult antisocial behavior (AAB) in the NCS-R were ODD (8.5 %), CD (9.5 %), ADHD (8.1 %), ALC (7.8 %), DRGMJ (5.1 %), and AAB (3.6 %).

Analysis Plan

To evaluate the fit of several comparison dimensional-categorical models, we followed the guidance of Masyn and colleagues (2010), who provide a framework for building and comparing hybrid models along a dimensional-categorical spectrum. Fully dimensional models (e.g., factor analysis, item response theory) assume that all common variability in the observed indicators can be explained by one or more continuous latent factors. A fully dimensional model is analogous to a continuous measure of externalizing behavior (e.g., total score on an ADHD checklist). Fully categorical models (e.g., latent class models) assume that shared variability in the observed indicators can be explained by one or more distinct latent classes that are homogenous with respect to the profiles of observed indicators. A fully categorical model could be compared to a classification system where individuals may be divided into mutually exclusive groups (e.g., hyperactive subtype versus impulsive subtype of ADHD). The factor mixture model (FMM) is a hybrid dimensional-categorical model that characterizes variability in the observed indicators as a categorical latent variable in which each latent class is defined by a continuous latent factor, thus allowing dimensionality within classes. The FMM is analogous to having dimensional symptoms (e.g., continuous levels of hyperactivity and impulsivity within hyperactive and impulsive subtypes). Krueger and colleagues (Krueger et al. 2005; Markon and Krueger 2005) have focused on a restricted form of the FMM in their prior studies (called a non-parametric factor model by Masyn et al. 2010), whereby the distribution of the observed indicators are estimated using a finite number of latent classes. The variation in observed indicators is thus defined by a mean level of the latent factor and it is assumed that there is no variability in dimensionality within the latent classes. For more information about factor mixture model specification, the interested reader is referred to numerous technical references and simulation studies that have examined the behavior of factor mixture models under various conditions (Lubke and Muthén 2007; Lubke and Neale 2006, 2008; Lubke and Spies 2008; Masyn et al. 2010).

In the current study, four models along the dimensional-categorical spectrum were estimated: (a) continuous latent factor models (fully dimensional), (b) non-parametric factor models (hybrid dimensional-categorical, defined by multiple latent classes without dimensionality within classes), (c) factor mixture models (hybrid dimensional-categorical, defined by multiple latent classes with dimensionality within classes), and (d) latent class models (fully categorical). Substantive and empirical decision rules were used to evaluate the factor structure all models (MacLachlan and Peel 2000; Nylund et al. 2007). Specifically, latent factor models with a varying number of factors were considered a good fit to the observed data based on non-significant χ^2 values, the Root Mean Square Error of Approximation (RMSEA) less than 0.06 (Browne and Cudeck 1993), and the Comparative Fit Index (CFI; Bentler 1990) and Tucker-Lewis Index (TLI; Tucker and Lewis 1973) greater than 0.95 (Hu and Bentler 1999). For models with latent classes, the Lo et al. (2001) likelihood ratio test (LRT) was used to test the number of classes, with a significant p value indicating that $k-1$ classes should be rejected in favor of at least k classes (Lo et al. 2001; Nylund et al. 2007).

For the factor models, we extended Krueger's (Krueger et al. 2005; Markon and Krueger 2005) work by estimating multidimensional factor models. Specifically, we tested the models shown in Fig. 1: a one-factor model (model 1), four alternative two-factor models (models 2a–2d), and a three-factor model (model 3). Models 1, 2a, 2b, and 3 were direct replications of the models from Farmer and colleagues (2009). Models 2c and 2d were included in the current study to replicate the analyses of Tuvblad and colleagues (2009). For model 2d, the decision to estimate the cross-loadings of AAB on both factors came from findings that AAB and SUD tend to load on a single genetic factor (e.g., Kendler et al. 2011) and that AAB also tends to be highly associated with ADHD, ODD, and CD (e.g., Fischer et al. 2002).

Given that models 2a and 2c were nested within model 2d by restricting AAB to load with either ADHD, ODD, and CD (model 2a) or SUD (model 2c), we were able to evaluate whether removing the cross-loading of AAB resulted in a significant decrement in model fit. Model comparisons between 2a and 2d, as well as 2c and 2d, were conducted using a χ^2 difference test, where a significant difference would indicate models 2a or 2c fit significantly worse than the less restrictive model (2d).

Results from the latent factor and latent class models were then used to guide specification of the non-parametric factor models and the FMMs. We tested one non-parametric factor model and three different FMMs. The non-parametric factor model assumed latent classes were located on a continuum with the variance of the latent factor within each class set to zero. In FMM-1, we constrained the factor loadings and thresholds to be equivalent across classes and only allowed the means and covariance of the factors to vary across classes. In FMM-2, we constrained the thresholds to be equivalent and freed the factor loadings, means, and covariance of the factors to vary across classes. In FMM-3, we allowed thresholds, factor loadings, factor means, and the covariance matrix to be freed across classes.

The optimal model from the selection of alternative models along the dimensional-categorical spectrum was selected using the Bayesian Information Criterion (BIC; Schwarz 1978) and the LRT. Simulation studies to evaluate the performance of FMMs, in comparison to continuous latent factor models and categorical latent class models, under various conditions and model specifications have found that the LRT (Lubke and Muthén 2007) and BIC (Lubke and Neale 2006, 2008) were most likely to identify the correct model when comparing the fit of several different models. The BIC has typically been used to make model selection decisions in applications of FMMs (Lubke et al. 2009; Lubke and Spies 2008; Walton et al. 2011). For the current study, we used the LRT to compare across mixture models within the same group of models (e.g., two-class versus three-class latent class model), whereby a non-significant LRT would indicate that a model with fewer classes would be selected. We used the BIC to compare across different model types (e.g., FMM vs. continuous latent factor), with lower values of BIC considered better-fitting models.

All models, described below, were estimated in Mplus version 6.1 (Muthén and Muthén 2010). To compare the fit of the latent factor models using standard model fit indices we estimated the latent factor model parameters for the diagnostic items using the weighted least squares estimator with means and variances adjusted procedure (WLSMV). WLSMV was chosen for initial model testing to replicate the analyses by Farmer and colleagues (2009). Second, to compare the fit of models across the dimensional-categorical spectrum, we estimated parameters using a weighted maximum likelihood (ML) function with all standard errors computed using a sandwich estimator. ML was chosen for comparing the dimensional-categorical models because (a) it was used by Krueger and colleagues (2005), (b) it was recommended by Masyn and colleagues (2010), and (c) because ML is a preferred

estimation method when some data are missing, assuming data are missing at random (Schafer and Graham 2002).

Results

The tetrachoric correlations between the ADHD, CD, and ODD diagnoses all exceeded 0.50 and the correlations between ALC, MJ and DRG diagnoses exceeded 0.47. AAB was most strongly associated with MJ ($r = 0.62$) and ODD ($r = 0.55$) diagnoses. The smallest correlations were between diagnoses of ALC and ADHD ($r = 0.001$), ALC and CD ($r = 0.10$), and DRG and ODD ($r = 0.13$).

Model Fit

Latent Factor Models—Model comparisons for the fully dimensional continuous latent factor models in the Fast Track sample are provided in Table 1. The RMSEA was below 0.06 for all models; however, the other indices of model fit suggested the one-factor model and model 2b (see Fig. 1), which delineated ADHD/ODD as indicators of one factor and AAB, CD, and SUD as indicators of a second factor, did not provide a good fit to the observed data. Model 2d, which allowed AAB to cross-load on an attention-deficit and disruptive behaviors (ADHD, CD, ODD) factor and a SUD factor, was selected as the best-fitting model as indicated by non-significant $\chi^2(12) = 12.29$ ($p = 0.42$), CFI/TLI = 0.99, and RMSEA = 0.006. The BIC for each of the factor models, provided in Table 2, also indicated model 2d provided the best fit to the observed data. Models 2a and 2c, and the three-factor model also provided a good fit to the observed data. Models 2a and 2c were similar to 2d in that all three models distinguished ADHD, ODD, and CD from SUD, yet χ^2 difference testing indicated that removing the cross-loading of AAB and restricting AAB to only load with either ADHD, ODD, and CD (model 2a) or SUD (model 2c) led to a significant decrement in model fit for both models (Model 2a: $\Delta\chi^2(1) = 8.19$, $p = 0.004$; Model 2c: $\Delta\chi^2(1) = 6.66$, $p = 0.009$). Model 3 could be considered nearly equivalent to model 2a, because the first two factors of the three-factor model 3 were correlated at 0.99.

Latent Class Models—Results from the two- and three-class fully categorical latent class models are presented in Table 2. The Lo-Mendell-Rubin likelihood ratio test (LRT) indicated that the two-class model fit significantly better than a one-class model and that a three-class model did not fit significantly better than a two-class model. The BIC also identified the two-class model as the best-fitting model.

Non-Parametric Factor Models—As seen in Table 2, the two-factor model with three values of the latent class variable provided the best fit based on BIC. The three-value model of the diagnostic items was rejected in favor of a two-value model by the LRT.

Factor Mixture Models—The LRT rejected two- and three-class FMMs in favor of a one-class model. The BIC identified a one-factor, two-class model as the best-fitting model for FMM-2 and FMM-3, but not for FMM-1. In all models there were many parameter estimates outside of acceptable ranges, small standardized loadings, and models explained nearly zero variance in some items.

Comparisons Across Models

As seen in Table 2, the fully dimensional latent factor models provided the best fit to the data, as indicated by lower BIC in comparison to the latent class and hybrid models. The LRT estimates for the factor mixture models, which allowed for dimensionality within class, indicated that the two-class factor mixture models were rejected in favor of one-class models, providing further evidence that the dimensional models provided a better fit to the

data than a categorical model. Among latent factor models the BIC was lowest for model 2d, which is consistent with results from Table 1.

Interpretation of Estimates from Model 2d and Alternative Factor Models

Among all estimated models, model 2d (i.e., the two-factor continuous confirmatory factor model with one factor indicated by ADHD, ODD, CD, and AAB and the second factor indicated by AAB and SUD) was selected as the best-fitting model. Because the latent factor models were estimated with two types of estimators (i.e., weighted least squares with means and variances adjusted (WLSMV) and maximum likelihood with robust standard errors (MLR)), we examined the factor loadings obtained from both estimation methods and found very similar results across estimation methods. ODD had the highest standardized loading (0.76 for MLR, 0.77 for WLSMV) on the disruptive/antisocial behavior factor and marijuana abuse/dependence had the highest loading on the antisocial/substance use factor (0.95 for MLR; 0.96 for WLSMV). Also, AAB had the lowest loadings on both factors with standardized loadings of 0.45 (MLR) and 0.45 (WLSMV) on the disruptive/antisocial behavior factor and 0.44 (MLR) and 0.46 (WLSMV) on the antisocial/substance use factor. Item parameter estimates and information curves for model 2d are available from the first author.

A second interpretation of model 2d could be that AAB was not a good indicator of either factor. Secondary analyses of model 2d without AAB also provided an excellent fit to the observed data ($\chi^2(8) = 10.76$ ($p = 0.22$), CFI/TLI = 0.99/0.97, and RMSEA = 0.022), suggesting the multidimensionality of externalizing disorders can be characterized by a disruptive behavior (ODD, ADHD, and CD) factor and a substance use factor without the inclusion of AAB. Another alternative model could be estimated with a single higher-order factor indicated by the disruptive behavior factor (ODD, ADHD, and CD), the substance use factor, and AAB. This model also provided a reasonable fit to the data ($\chi^2(13) = 26.05$ ($p = 0.02$), CFI/TLI = 0.95/0.92, and RMSEA = 0.037) with all factor loadings exceeding 0.65. Thus, AAB was strongly associated with both disruptive behaviors and SUD, but AAB is not necessary to define the structure of externalizing behavior delineated by disruptive behaviors and substance use.

To examine the covariation between disruptive behaviors, substance use, and AAB, without including AAB as an indicator, we conducted additional analyses with AAB symptoms regressed on the disruptive behaviors and substance use factors. Squared semi-partial correlation coefficients indicated that 12 % of the variance in AAB was explained by shared variance between the disruptive behaviors and substance use factors, 5 % of the variance in AAB was uniquely explained by disruptive behaviors, and 8 % of the variance in AAB was uniquely explained by substance use. These results highlight that while AAB was correlated with CD, ODD, ADHD, and SUD, there was also a good deal of variance (75 %) that was not explained by these externalizing disorders. Likewise, only 31% of the variance in the substance use factor was shared by the disruptive behaviors factor, supporting the notion that a single externalizing dimension may not be sufficient to explain covariation between these disorders.

Replication of Model 2d and Alternative Models in the NCS-R Sample

A final goal of the current study was to determine whether the confirmatory factor model of externalizing disorders selected in the current study, model 2d, would replicate in a different sample. To accomplish this goal we estimated the series of confirmatory factor models described above (see Fig. 1) using data from the NCS-R. Results from the confirmatory factor models of the NCS-R data are provided in Table 3. Consistent with the Fast Track sample analyses, model 2d provided an excellent fit to the data based on all indicators.

Models 2a, 2c and the three-factor model also provided an excellent fit. Nested model testing indicated that removing the cross-loading in model 2a and 2c did result in a significant decrement in model fit compared to model 2d (model 2a: $\Delta\chi^2(1) = 9.76, p = 0.002$; model 2c: $\Delta\chi^2(1) = 51.30, p < 0.0001$). As in the Fast Track sample, the three-factor model was very similar to model 2a with a 0.76 correlation between the first two factors of the three-factor model, suggesting that model 2a could be selected as a more parsimonious alternative model.

There were some differences in the factor loadings for the NCS-R model 2d (right of the path in Fig. 2), in comparison to the Fast Track sample results (left of the path in Fig. 2). In the NCS-R two-factor model, CD had the highest loading on the disruptive behavior factor (WLSMV standardized loading = 0.91), and drug abuse/dependence had the highest loading on the antisocial/substance use factor (WLSMV standardized loading = 0.94). Again, AAB had the lowest loadings on both factors.

The secondary analyses of model 2d without AAB also provided an excellent fit to the observed data in the NCS-R sample ($\chi^2(4) = 19.27 (p = 0.0007)$, CFI/TLI = 0.99/0.98, and RMSEA = 0.028), suggesting that AAB could be excluded from the model without significant decrement of model interpretation or fit. The second alternative model with a single higher-order factor indicated by the disruptive behavior factor, the substance use factor, and AAB also provided a reasonable fit to the data in the NCS-R sample ($\chi^2(8) = 23.62 (p = 0.002)$, CFI/TLI = 0.99/0.99, and RMSEA = 0.018) with all factor loadings exceeding 0.57 in the NCS-R sample. Thus, AAB could be considered strongly associated with both disruptive behaviors and SUD, but AAB is not necessary to define the structure of externalizing behavior delineated by disruptive behaviors and substance use.

Discussion

The goal of the current study was to replicate and extend recent evaluations of the externalizing spectrum, which have advocated that a single dimensional externalizing latent factor can explain the covariation in externalizing behavior disorders (Krueger et al. 2005; Markon and Krueger 2005). The results from the present research provide additional support for a dimensional conceptualization of externalizing disorders. However, consistent with recent work by Farmer and colleagues (2009), findings from the current study suggested that a single latent factor was not sufficient to explain the covariation in externalizing disorders across two separate samples, particularly when childhood-onset disorders are considered. Results from our study provide support for a two-factor model of externalizing psychopathology characterized by a factor of hyperactivity/impulsivity, oppositionality, and conduct disorder/antisocial behaviors, that is correlated with a factor of antisocial and substance use disorders.

Antisocial personality characteristics could be included as a symptom indicator of both hyperactive/oppositional behavior disorders and SUDs. However, covariation between ADHD, ODD, and CD were distinguished from covariation in SUDs whether or not AAB was included in the model. In fact, all of the models that incorporated ADHD, ODD, and CD diagnoses as factors separate from SUD (models 2fa, 2fc, 2fd, 3f) provided an excellent fit to the data based on all indicators in both the Fast Track and NCS-R samples, and all of these models distinguished ADHD, ODD, and CD from SUD. Models that provided the worst fit to the data across both samples (model 2fb, 1f) included SUD with CD and AAB (model 2fb) or all externalizing disorders (model 1f) as indicators of a single latent factor.

Comparison of Current Results to Prior Studies of the Externalizing Spectrum

Findings from the current study were consistent with previous studies that have concluded dimensional models of externalizing psychopathology provide a better fit than categorical models (Krueger et al. 2005; Markon and Krueger 2005). The current study extends this work by demonstrating that ODD and ADHD can also be considered part of a dimensional model of externalizing disorders, which is consistent with descriptions of the externalizing spectrum in a developmental context (Tackett 2010). The prior work by Krueger (Krueger et al. 2005; Markon and Krueger 2005) has found CD to load on a single externalizing factor with ASPD and SUDs, but that prior work has not considered other externalizing behaviors that tend to precede ASPD and SUDs developmentally. Given the stability of externalizing behavior symptoms over time, as well as the possibility of shared etiologies (e.g., trait impulsivity; Beauchaine et al. 2010), it is important to consider the full spectrum of externalizing behaviors in models of the “externalizing spectrum.”

Moreover, our findings suggest that when ADHD and ODD are included in the broader externalizing spectrum model, they load onto a factor with CD. These findings are consistent with research that has supported Achenbach’s (1966) hierarchical classification of internalizing and externalizing behavior disorders and are also consistent with the documentation of high rates of comorbidity among ADHD, CD, and ODD (e.g., Burt et al. 2005), recent genetic analyses (Tuvblad et al. 2009), and a biological vulnerability-by-environmental risk model of externalizing disorders (Kendler et al. 2011). Importantly, the fact that ADHD and ODD loaded on a single factor with CD in the current study does not imply that ADHD, ODD, and CD should be “lumped” together. The current findings imply that there is shared variance among the disorders but important distinctions between ADHD, ODD, and CD have been noted (Bezdjian et al. 2011).

The second externalizing factor was represented by covariance between AAB and SUD. Although highly correlated with the disruptive/antisocial behaviors factor ($r = 0.50$), these findings suggest that covariation between AAB and SUD can be modeled separately from the externalizing behaviors commonly observed in childhood. Thus, SUD may be influenced by early externalizing behaviors, but it may not be accurate to describe SUD as an adolescent/adult manifestation of early externalizing behaviors. Multiple other factors, such as parenting, peer networks, physiological sensitivity to the effects of alcohol and drugs (Chassin et al. 2002) and other developmental pathways [e.g., depression or anxiety (Kaplow et al. 2001)] have been connected to risk for SUDs and it may be that SUD did not load onto the same factor as CD/ODD or ADHD because of variation in how individuals follow pathways from early externalizing disorders to later SUD. It is also the case that different trajectories of substance use onset and persistence have been described (e.g., Jackson et al. 2008) and different trajectories of substance use may be uniquely related to different trajectories of disruptive behaviors (Marti et al. 2010; Wu et al. 2010).

Limitations

The primary limitation of the current study was the measurement of psychiatric diagnoses in the Fast Track and NCS-R samples. Consistent with many longitudinal studies, the Fast Track study used different measurement procedures across time, with both the decision rules for the diagnoses and the reporter changing over time. The fact that some individuals (African Americans, individuals with a lifetime diagnosis of ADHD or ODD) were more likely to be missing data on the AAB assessment is also a limitation. The replication of models in the NCS-R data, which was based on a single assessment device and single reporter, helps mitigate some of these concerns. However, an important limitation of the NCS-R data was that diagnoses were based on recall of childhood behavior problems, which could reduce the accuracy of diagnoses (Barkley et al. 2002). Likewise, the AAB variable

derived from the Fast Track and NCS-R data was not based on full DSM-IV symptoms of ASPD. More importantly, across both samples the dimensionality of the externalizing spectrum was evaluated using diagnosis-driven measurement with presence or absence of certain symptoms as indicators of diagnosis. It may have been preferable to use scales or behavioral measures since these measures might be more sensitive to continuous differences in externalizing behaviors. The Fast Track models described in this study were successfully replicated with continuous criterion counts (results available from the first author), but the criterion counts are still based on the DSM system. Consistent with prior categorical-dimensional examinations of the externalizing spectrum (e.g., Krueger et al. 2005; Markon and Krueger 2005), we did not evaluate the externalizing spectrum model over time. Specifically, measurement limitations in both the Fast Track data (potential for non-invariance over time) and NCS-R data (only life-time diagnoses) led us to rely on lifetime diagnoses and we did not examine developmental changes in externalizing symptoms. Importantly, recent work by our research team (King et al. 2012) found support for our dimensional model of the externalizing spectrum measured prospectively in the Fast Track data from childhood to young adulthood. Results indicated clear autoregressive pathways whereby covariation among externalizing symptoms in earlier years (starting in kindergarten) was prospectively related to the covariation in later years through age 20.

A third limitation of the current study was that countless alternative models could have provided an equal or better fit to the data and a well-fitting model is not necessarily free of misspecifications (Tomarken and Waller 2003). Replication of the factor structure in the NCS-R data provides evidence in support of the model identified in the Fast Track data. Yet, alternative modeling approaches (e.g., bifactor models) might also provide a useful characterization of the externalizing spectrum.

Implications for DSM-5

It was proposed that DSM-5 include a section on the dimensionality of externalizing behavior disorders (Krueger and South 2009), incorporating CD, SUD, and perhaps borderline personality disorder and ADHD, into a single externalizing cluster. However this proposal was based on studies that only considered a single dimension of externalizing behavior, whereas our results suggest that a multidimensional approach might be necessary. In contrast to previous proposals put forward by Krueger and his colleagues (2005), our results indicate that ADHD, CD and ODD should be considered as relatively distinct from (though related to) SUD, with AAB sharing commonalities with both. Such a multidimensional perspective on externalizing disorders implies that there may be different etiological or developmental pathways to and from the disruptive and antisocial factor, as compared to the antisocial and SUD factor. Thus, considering CD and SUD to be part of the same unidimensional factor (see Krueger et al. 2005; Markon and Krueger 2005) may obscure these distinct pathways.

Considering distinct but related factors for CD and SUD would acknowledge that SUD can also be identified as an indicator of internalizing disorders (e.g., Kendler et al. 2011). Similarly, recent work has found that ODD is multidimensional with multiple etiologies and distinct developmental pathways that may differ by gender (Burke and Loeber 2010; Rowe et al. 2010). ODD can be characterized by both affective and behavioral symptoms, with the affective symptoms being predictive of later depression (“internalizing”) and the behavioral symptoms being associated with CD and aggressive behaviors (“externalizing”; Burke and Loeber 2010). These findings have led for a call to maintain distinctions between ODD and CD in DSM-5 (Rowe et al. 2010). Results from the current study should be replicated prior to making any statements about changes to the DSM-5, particularly given the conceptual and practical issues involved in changing a diagnostic system.

Conclusions and Future Directions

In order to gain a better understanding of the structure of psychopathology in a way that will significantly advance research and treatment, quantitative investigations need to be extended beyond the goals of the current study, which focused on symptom/disorder overlap, and towards the goal of understanding the broader etiology and progression of psychopathology over time. Future research on the risk and protective factors and identification of subordinate factors within the externalizing factors could also provide valuable information on how externalizing psychopathology may be modified over time. For example, research has identified subtypes of CD (Tackett et al. 2003), ADHD (Todd et al. 2001), ODD (Rowe et al. 2010), and SUD (Winters et al. 2008) and these different subtypes might respond differently to specific treatments or be identified by distinct phenotypes. Estimating subordinate factors that cut across diagnostic categories could help characterize individuals at the level of behavior, or even better at the level of common vulnerabilities (Beauchaine et al. 2010), rather than at the level of disorder. Such theoretically-informed, quantitative investigations that take a multidimensional approach to measuring externalizing psychopathology may provide a better approximation of externalizing symptomatology and will inform future research on the etiology, prevention, and treatment of externalizing behavior problems.

Acknowledgments

This work was supported by National Institute of Mental Health (NIMH) grants R18 MH48043, R18 MH50951, R18 MH50952, R18 MH50953, and by National Institute on Drug Abuse (NIDA) grant 1 RC1 DA028248-01. The Center for Substance Abuse Prevention and NIDA also provided support for Fast Track through a memorandum of agreement with the NIMH. This work was also supported in part by Department of Education grant S184U30002, NIMH grants K05MH00797 and K05MH01027, and NIDA grants DA16903, DA015226 and DA017589.

References

- Achenbach TM. The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs*. 1966; 80:1–37. [PubMed: 5968338]
- Achenbach, TM. *Manual for the child behavior checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry; 1991.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th ed., text rev.. Washington, DC: Author; 2000.
- American Psychiatric Association. DSM-5 development. 2010. Cross-cutting dimensional assessment in DSM-5. Retrieved from <http://www.dsm5.org/ProposedRevisions/Pages/CrossCuttingDimensionalAssessmentinDSM-5.aspx>.
- Barkley R, Fischer M, Smallish L, Fletcher K. The persistence of attention deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology*. 2002; 111:279–289. [PubMed: 12003449]
- Beauchaine TP, Hinshaw SP, Pang KL. Comorbidity of attention-deficit/hyperactivity disorder and early-onset conduct disorder: Biological, environmental, and developmental mechanisms. *Clinical Psychology: Science and Practice*. 2010; 17:327–336.
- Bentler P. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107:238–246. [PubMed: 2320703]
- Bezdjian S, Krueger RF, Derringer J, Malone S, McGue M, Iacono WG. The structure of DSM-IV ADHD, ODD, and CD criteria in adolescent boys: A hierarchical approach. *Psychiatry Research*. 2011; 188:411–421. [PubMed: 21470694]
- Brown TA, Barlow DH. Categorical versus dimensional classification of mental disorders in *DSM-V* and beyond. *Journal of Abnormal Psychology*. 2005; 114:551–556. [PubMed: 16351377]
- Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Beverly Hills, CA: Sage; 1993. p. 136-162.

- Burt SA, McGue M, Krueger RF, Iacono W. How are parent-child conflict and childhood externalizing symptoms related over time? results from a genetically informative cross-lagged study. *Development and Psychopathology*. 2005; 17:145-165. [PubMed: 15971764]
- Burke J, Loeber R. Oppositional defiant disorder and the explanation of the comorbidity between behavioral disorders and depression. *Clinical Psychology: Science and Practice*. 2010; 17:319-326.
- Chassin L, Pitts SC, Prost J. Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: Predictors and substance abuse outcomes. *Journal of Consulting and Clinical Psychology*. 2002; 70:67-78. [PubMed: 11860058]
- Conduct Problems Prevention Research Group. A developmental and clinical model for the prevention of conduct disorders: The FAST Track program. *Development and Psychopathology*. 1992; 4:509-527.
- Conduct Problems Prevention Research Group. Initial impact of the Fast Track prevention trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology*. 1999; 67:631-647. [PubMed: 10535230]
- Conduct Problems Prevention Research Group. Merging universal and indicated prevention programs: The Fast Track model. *Addictive Behaviors*. 2000; 25:913-927. [PubMed: 11125779]
- Farmer RF, Seeley JR, Kosty DB, Lewinsohn PM. Refinements in the hierarchical structure of externalizing psychiatric disorders: Patterns of lifetime liability from mid-adolescence through early adulthood. *Journal of Abnormal Psychology*. 2009; 118:699-710. [PubMed: 19899840]
- First MB. Defining mental disorder in DSM-V. *Psychological Medicine*. 2010; 40:1779-1782. [PubMed: 20102663]
- Fischer M, Barkley RA, Smallish L, Fletcher K. Young adult follow-up of hyperactive children: Self-reported psychiatric disorders, comorbidity, and the role of childhood conduct problems and teen CD. *Journal of Abnormal Child Psychology*. 2002; 30:463-475. [PubMed: 12403150]
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1-55.
- Jablensky A. A meta-commentary on the proposal for a meta-structure for DSM-V and ICD-11. *Psychological Medicine*. 2009; 39:2099-2103. [PubMed: 19796433]
- Jackson KM, Sher KJ, Schulenberg JE. Conjoint developmental trajectories of young adult substance use. *Alcoholism, Clinical and Experimental Research*. 2008; 32:723-737.
- Kamphuis JH, Noordhof A. On categorical diagnoses in DSM-V: Cutting dimensions at useful points? *Psychological Assessment*. 2009; 21:294-301. [PubMed: 19719342]
- Kaplow JB, Curran PJ, Angold A, Costello EJ. The prospective relation between dimensions of anxiety and the initiation of adolescent alcohol use. *Journal of Clinical Child Psychology*. 2001; 30:316-326. [PubMed: 11501249]
- Kendler KS, Aggen SH, Knudsen GP, Roysmb E, Neale MC, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *The American Journal of Psychiatry*. 2011; 168:29-39. [PubMed: 20952461]
- Kessler RC, Ustun TB. The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*. 2004; 13:93-121. [PubMed: 15297906]
- Kessler RC, Berglund P, Chiu WT, Demler O, Heeringa S, Hiripi E, et al. The US National Comorbidity Survey Replication (NCS-R): Design and field procedures. *International Journal of Methods in Psychiatric Research*. 2004; 13:69-92. [PubMed: 15297905]
- King, KM.; Luk, JW.; Wu, J.; Witkiewitz, K.; Racz, SJ.; McMahon, RJ. the Conduct Problems Prevention Research Group. The co-occurrence of externalizing behaviors during childhood: Factor structure and invariance over time. 2012. Manuscript in preparation.
- Krueger RF, Bezdjian S. Enhancing research and treatment of mental disorders with dimensional concepts: Toward DSM-V and ICD-11. *World Psychiatry*. 2009; 8:3-6. [PubMed: 19293948]
- Krueger RF, Hicks BM, Patrick CJ, Carlson SR, Iacono WG, McGue M. Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology*. 2002; 111:411-424. [PubMed: 12150417]

- Krueger RF, Markon KE, Patrick CJ, Benning SD, Kramer MD. Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology*. 2007; 116:645–666. [PubMed: 18020714]
- Krueger RF, Markon KE, Patrick CJ, Iacono WG. Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implications for DSM-V. *Journal of Abnormal Psychology*. 2005; 114:537–550. [PubMed: 16351376]
- Krueger RF, South SC. Externalizing disorders: Cluster 5 of the proposed meta-structure for DSM-V and ICD-11. *Psychological Medicine*. 2009; 39:2061–2070. [PubMed: 19796431]
- Lahey BB, Rathouz PJ, Applegate B, Van Hulle CA, Garriock HA, Urbano RC, et al. Testing structural models of DSM-IV symptoms of common forms of child and adolescent psychopathology. *Journal of Abnormal Child Psychology*. 2008; 36:187–206. [PubMed: 17912624]
- Lahey BB, Van Hulle CA, Singh AL, Waldman ID, Rathouz PJ. Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Archives of General Psychiatry*. 2011; 68:181–189. [PubMed: 21300945]
- Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. *Biometrika*. 2001; 88:767–778.
- Loeber R, Burke JD. Developmental pathways in juvenile externalizing and internalizing problems. *Journal of Research on Adolescence*. 2011; 21:34–46. [PubMed: 22468115]
- Lubke GH, Hudziak JJ, Derks EM, van Bijsterveldt TC, Boomsma DI. Maternal ratings of attention problems in ADHD: Evidence for the existence of a continuum. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2009; 48:1085–1093. [PubMed: 19797980]
- Lubke GH, Muthén BO. Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling*. 2007; 14:26–47.
- Lubke GH, Neale MC. Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood? *Multivariate Behavioral Research*. 2006; 41:499–532.
- Lubke GH, Neale MC. Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models. *Multivariate Behavioral Research*. 2008; 43:592–620. [PubMed: 20165736]
- Lubke, GH.; Spies, J. Choosing a ‘correct’ factor mixture model: Power, limitations, and graphical data exploration. In: Hancock, GR.; Samuelsen, KM., editors. *Advances in latent variable mixture models*. Charlotte, NC: Information Age Publishing; 2008. p. 343-362.
- MacLachlan, GJ.; Peel, D. *Finite mixture models*. New York, NY: Wiley; 2000.
- Markon KE, Chmielewski M, Miller CJ. The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*. 2011; 137:856–879. [PubMed: 21574681]
- Markon KE, Krueger RF. Categorical and continuous models of liability to externalizing disorders: A direct comparison in NESARC. *Archives of General Psychiatry*. 2005; 62:1352–1359. [PubMed: 16330723]
- Marti CN, Stice E, Springer DW. Substance use and abuse trajectories across adolescence: A latent trajectory analysis of a community-recruited sample of girls. *Journal of Adolescence*. 2010; 33:449–461. [PubMed: 19640578]
- Masyn K, Henderson CE, Greenbaum PE. Exploring the latent structures of psychological constructs in social development using the dimensional-categorical spectrum. *Social Development*. 2010; 19:470–493.
- Muthén, B.; Muthén, L. *Mplus user’s guide*. 6th ed.. Los Angeles, CA: Authors; 2010.
- Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation. *Structural Equation Modeling*. 2007; 14:535–569.
- Rowe R, Costello EJ, Angold A, Copeland WE, Maughan B. Developmental pathways in oppositional defiant disorder and conduct disorder. *Journal of Abnormal Psychology*. 2010; 119:726–738. [PubMed: 21090876]
- Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychological Methods*. 2002; 7:147–177. [PubMed: 12090408]

- Schwarz G. Estimating the dimension of a model. *The Annals of Statistics*. 1978; 6:461–464.
- Shaffer, D.; Fisher, P. NIMH-diagnostic interview schedule for children: Child informant. New York: New York State Psychiatric Institute; 1997.
- Shaffer D, Fisher P, Lucas C, Dulcan M, Schwab-Stone M. NIMH diagnostic interview schedule for children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2000; 39:28–38. [PubMed: 10638065]
- Shaffer, D.; Fisher, P.; Lucas, C.; Comer, J. Scoring manual: Diagnostic interview schedule for children (DISC-IV). New York, NY: Columbia University; 2003.
- Tackett JL. Toward an externalizing spectrum in DSM-V: Incorporating developmental concerns. *Child Development Perspectives*. 2010; 4:161–167.
- Tackett JL, Krueger RF, Sawyer MG, Graetz BW. Subfactors of DSM-IV conduct disorder: Evidence and connections with syndromes from the child behavior checklist. *Journal of Abnormal Child Psychology*. 2003; 31:647–654. [PubMed: 14658744]
- Todd RD, Rasmussen ER, Neuman RJ, Reich W, Hudziak JJ, Bucholz KK, et al. Familiarity and heritability of subtypes of attention deficit hyperactivity disorder in a population sample of adolescent female twins. *The American Journal of Psychiatry*. 2001; 158:1891–1898. [PubMed: 11691697]
- Tomarken AJ, Waller NG. Potential problems with “well fitting” models. *Journal of Abnormal Psychology*. 2003; 112:578–598. [PubMed: 14674870]
- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973; 38:1–10.
- Tuvblad C, Zheng M, Raine A, Baker LA. A common genetic factor explains the covariation among ADHD ODD and CD symptoms in 9–10 year old boys and girls. *Journal of Abnormal Child Psychology*. 2009; 37:153–167. [PubMed: 19015975]
- Verona E, Javdani S, Sprague J. Comparing factor structures of adolescent psychopathology. *Psychological Assessment*. 2011; 23:545–551. [PubMed: 21280958]
- Walton KE, Ormel J, Krueger RF. The dimensional nature of externalizing behaviors in adolescence: Evidence from a direct comparison of categorical, dimensional, and hybrid models. *Journal of Abnormal Child Psychology*. 2011; 39:553–561. [PubMed: 21210204]
- Werthamer-Larsson L, Kellam S, Wheeler L. Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology*. 1991; 19:585–602. [PubMed: 1755437]
- Widiger TA, Samuel DB. Diagnostic categories or dimensions: A question for DSM-V. *Journal of Abnormal Psychology*. 2005; 114:494–504. [PubMed: 16351373]
- Winters KC, Stinchfield RD, Latimer WW, Stone A. Internalizing and externalizing behaviors and their association with the treatment of adolescents with substance abuse disorder. *Journal of Substance Abuse Treatment*. 2008; 35:269–278. [PubMed: 18328664]
- World Health Organization. International statistical classification of disease and related health problems, tenth revision (ICD-10). Geneva: Author; 1992.
- Wu J, Witkiewitz K, McMahon RJ, Dodge KA. the Conduct Problems Prevention Research Group. A parallel process growth mixture model of conduct problems and substance use with risky sexual behavior. *Drug and Alcohol Dependence*. 2010; 111:207–214. [PubMed: 20558013]

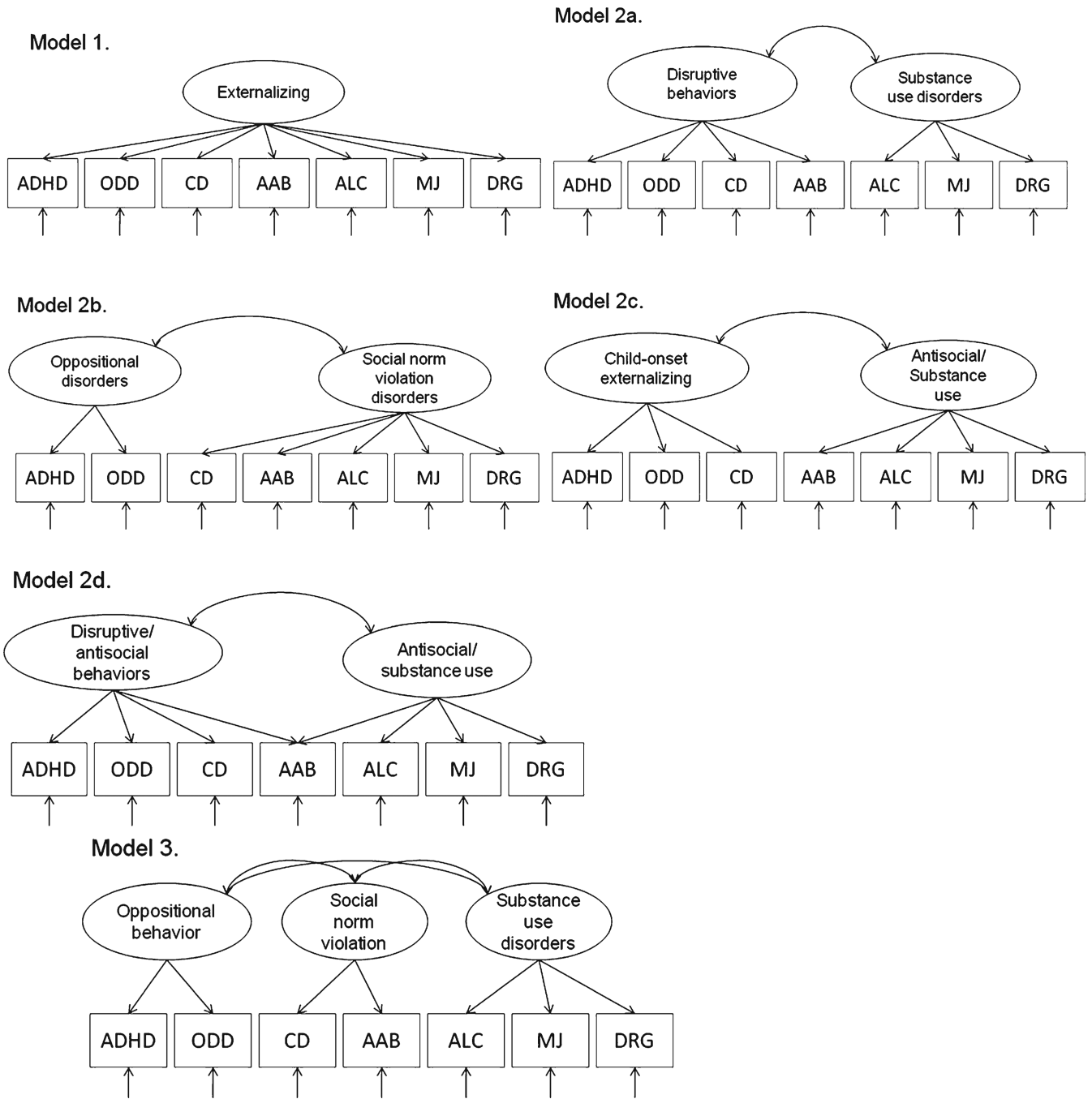


Fig. 1. Summary of confirmatory factor models. ADHD = Attention-deficit hyperactivity disorder; ODD = Oppositional defiant disorder; CD = Conduct disorder; AAB = Adult antisocial behavior; ALC = Alcohol abuse or dependence; MJ = Marijuana abuse or dependence; DRG = Other substance abuse or dependence

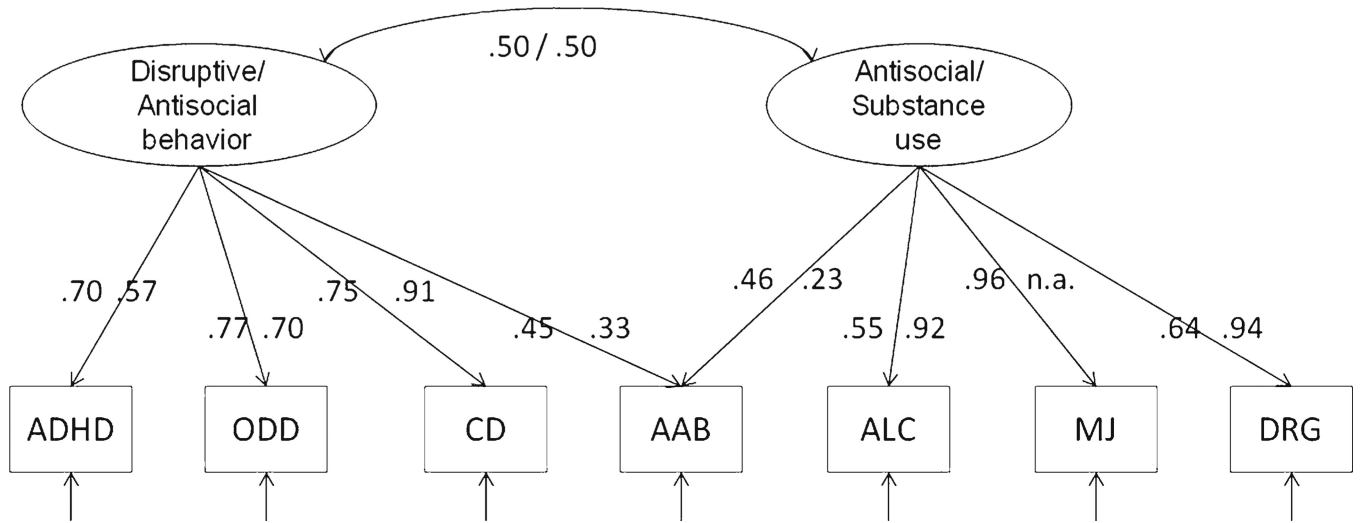


Fig. 2. Model 2d factor loadings (estimated using WLSMV) for the Fast Track sample (left side of each path) and the NCS-R sample (right side of each path). The factor loading for the MJ variable was not available (n.a.) for the NCS-R data because the NCS-R only included a measure of drug abuse or dependence and did not include a separate measure of marijuana abuse or dependence. AAB = Adult antisocial behavior; ALC = Alcohol abuse or dependence; MJ = Marijuana abuse or dependence; DRG = Other substance abuse or dependence

Table 1

Summary of model fit indices for the confirmatory factor models depicted in Fig. 1 for the fast track sample ($N=715$)

Psychiatric diagnosis models	χ^2 (df)	CFI	TLI	RMSEA	$\beta_{F1 \text{ with } F2}$	$\beta_{F1-F3/F2-F3}$
CFA 1 factor	37.25 (14), $p=0.001$	0.91	0.86	0.048	–	–
CFA 2fa: ADHD/ODD/CD/AAB vs. SUD	21.32 (13), $p=0.07$	0.97	0.95	0.030	0.62	–
CFA 2fb: ADHD/ODD vs. CD/AAB/SUD	34.27 (13), $p=0.001$	0.91	0.86	0.048	0.77	–
CFA 2fc: ADHD/ODD/CD vs. APSD/SUD	19.72 (13), $p=0.10$	0.97	0.96	0.030	0.65	–
CFA 2fd: ADHD/ODD/CD/AAB vs. AAB/SUD	12.29 (12), $p=0.42$	0.99	0.99	0.006	0.50	–
CFA 3f: ADHD/ODD vs. CD/AAB vs. SUD	16.09 (11), $p=0.14$	0.98	0.97	0.025	0.99	0.47/0.79

CFA Confirmatory factor analysis; ADHD Attention-deficit hyperactivity disorder; ODD Oppositional defiant disorder; CD Conduct disorder; AAB Adult antisocial behavior; SUD substance use disorders; β correlation between factors

Table 2
Model fit for the latent factor, latent class, and hybrid models for the fast track sample ($N=715$)

Model	Log-likelihood	# parameters	Entropy	BIC	LRT (p)
Latent factor models					
1-factor	-1313.76	14	1.00	2719.53	-
2-factor a	-1303.60	15	1.00	2705.78	-
2-factor b	-1310.60	15	1.00	2719.67	-
2-factor c	-1302.19	15	1.00	2702.97	-
2-factor d	-1298.26	16	1.00	2701.67	-
3-factor	-1300.74	17	1.00	2713.20	-
Latent class models					
2-class	-1317.87	15	0.70	2734.32	230.68 (0.0001)
3-class	-1299.07	23	0.81	2749.31	36.89 (0.51)
Non-parametric factor analysis					
1-factor, 2-value	-1317.87	15	0.70	2734.32	230.68 (0.0001)
1-factor, 3-value	-1311.04	17	0.67	2733.81	12.69 (0.79)
2-factor 2d, 2-value	-1317.87	16	0.70	2740.90	231.16 (0.0001)
2-factor 2d, 3-value	+1299.74	19	0.83	2724.36	34.50 (0.66)
FMM-1: class invariant factor loadings and thresholds; class varying factor means, and covariance matrix					
1-factor, 2-class	-1311.76	17	0.28	2735.25	3.71 (1.00)
2-factor 2d, 2-class	-1294.00	22	0.16	2732.59	9.86 (0.68)
FMM-2: class invariant thresholds; class varying factor loadings, means, and covariance matrix					
1-factor, 2-class	-1293.39	25	0.59	2751.08	40.14 (0.45)
2-factor 2d, 2-class	-1285.47	32	0.55	2781.25	25.38 (1.00)
MM-3: class varying factor loadings, thresholds, means, and covariance matrix					
1-factor, 2-class	-1288.28	32	0.85	2786.87	50.51 (0.95)
2-factor 2d, 2-class	-1283.45	39	0.30	2823.22	29.40 (1.00)

Table 3

Summary of model fit indices for the replication of confirmatory factor models using the NCS-R data ($N=5,692$)

Model	χ^2 (df)	CFI	TLI	RMSEA	$\beta_{F1 \text{ with } F2}$	$\beta_{F1-F3/F2-F3}$
CFA 1 factor	138.27 (9), $p<0.001$	0.94	0.89	0.049	–	–
CFA 2fa: ADHD/ODD/CD/AAB vs. SUD	37.75 (8), $p<0.001$	0.99	0.97	0.025	0.56	–
CFA 2fb: ADHD/ODD vs. CD/AAB/SUD	114.83 (8), $p<0.001$	0.95	0.90	0.047	0.61	–
CFA 2fc: ADHD/ODD/CD vs. AAB/SUD	51.30 (8), $p<0.001$	0.98	0.96	0.030	0.57	–
CFA 2fd: ADHD/ODD/CD/AAB vs. AAB/SUD	26.89 (7), $p=0.0003$	0.99	0.98	0.022	0.50	–
CFA 3f: ADHD/ODD vs. CD/AAB vs. SUD	19.95 (6), $p=0.003$	0.99	0.99	0.020	0.76	0.35/0.63

CFA Confirmatory factor analysis; ADHD Attention-deficit hyperactivity disorder; ODD Oppositional defiant disorder; CD Conduct disorder; AAB Adult antisocial behavior; SUD substance use disorders; β correlation between factors