Published in final edited form as:

J Abnorm Psychol. 2009 November ; 118(4): 699. doi:10.1037/a0017205.

Refinements in the Hierarchical Structure of Externalizing Psychiatric Disorders: Patterns of Lifetime Liability from Mid-Adolescence through Early Adulthood

Richard F. Farmer, John R. Seeley, Derek B. Kosty, and Peter M. Lewinsohn
Oregon Research Institute

Abstract

Research on hierarchical modeling of psychopathology has frequently identified two higher-order latent factors, “internalizing” and “externalizing.” When based on the comorbidity of psychiatric diagnoses, the externalizing domain has usually been modeled as a single latent factor. Multivariate studies of externalizing symptom features, however, suggest multidimensionality. To address this apparent contradiction, confirmatory factor analytic methods and information-theoretic criteria were used to evaluate four theoretically plausible measurement models based on lifetime comorbidity patterns of seven putative externalizing disorders. Diagnostic information was collected at four assessment waves from an age-based cohort of 816 persons between the ages of 14 and 33. A two-factor model that distinguished oppositional behavior disorders (attention-deficit/hyperactivity disorder, oppositional defiant disorder) from social norm violation disorders (conduct disorder, adult antisocial behavior, alcohol use disorder, cannabis use disorder, hard drug use disorder) demonstrated consistently good fit and superior approximating abilities. Analyses of psychosocial outcomes measured at the last assessment wave supported the validity of this two-factor model. Implications of this research for the theoretical understanding of domain-related disorders and the organization of classification systems are discussed.

Keywords

Externalizing disorders; measurement model; substance use disorders; disruptive behavior disorders; attention-deficit/hyperactivity disorder

Non-random patterns of diagnostic comorbidity among some combinations of psychiatric disorders are common and likely meaningful. In the National Comorbidity Survey (NCS; Kessler et al., 1994), for example, 79% of persons diagnosed with a lifetime psychiatric disorder also met criteria for at least one other psychiatric disorder during their lifetime. Additionally, a comparatively small proportion of the NCS sample (14%) that reported a psychiatric history of three or more comorbid disorders accounted for 59% of all diagnosed lifetime disorders. Kessler et al. (2005a) and Kessler, Chiu, Demler, and Walters (2005b) reported similar findings based on lifetime and 12-month comorbidity estimates, respectively, from the NCS replication sample. Such patterns of comorbidity among psychiatric disorders highlight possible common etiological processes, genetic influences, or maintaining factors among subsets of disorders, and may also have implications for treatment selection and responsiveness to specific therapies (Krueger, 1999).

In several recent reports, confirmatory factor analytic (CFA) methods have been used to evaluate competing hierarchical models of psychiatric disorders based on concurrent, 12-month, or lifetime diagnostic comorbidity. A frequent assumption underlying this research is that the resultant measurement models reveal a “liability spectrum,” whereby certain psychiatric disorders are regarded as expressions of latent liabilities that, in turn, explain diagnostic comorbidity or the increased risk for spectrum-related disorders during one’s lifetime (Krueger & Markon, 2006). This research has produced remarkably consistent findings, revealing a hierarchical organizational structure defined by two higher-order latent factors, *internalizing* and *externalizing*.

The internalizing liability spectrum, which accounts for the substantial comorbidity among the mood and anxiety disorders (Watson, O’Hara, & Stuart, 2008), has been frequently bifurcated into two lower-level latent factors (Cox, Clara, & Enns, 2002; Krueger, 1999; Krueger & Markon, 2006; Slade & Watson, 2006; Vollebergh et al., 2001). One of these subfactors is usually labeled “distress” or “mood,” and typically defined by depressive disorders (i.e., major depressive disorder, dysthymia) and generalized anxiety disorder. The second subfactor, “fear” or “anxiety,” usually includes a variety of phobic conditions and panic disorder.

In contrast to the internalizing spectrum, less is known concerning the inclusiveness and optimal modeling of putative externalizing psychiatric disorders. Modeling studies of externalizing diagnostic categories (Krueger, 1999; Krueger, Caspi, Moffitt, & Silva, 1998; Krueger & Markon, 2006; Slade & Watson, 2006; Vollebergh et al., 2001; see also a similar genetic modeling study by Kendler, Prescott, Myers, & Neale, 2003) consistently report that a single latent factor best accounts for these disorders. In these studies, however, limited sets of externalizing disorders have been evaluated (≤ 4 diagnostic categories), and putative internalizing disorders were also included in the measurement models. Of these studies, Krueger and Markon’s (2006) meta-analytic study was the only one to test an alternative model that included more than one latent externalizing factor. They nonetheless reported that a 3-factor model that consisted of a single externalizing factor and two internalizing factors fit the data best.

Other studies have evaluated externalizing models based on symptom counts of psychiatric diagnostic categories (Krueger, Chentsova-Dutton, Markon, Goldberg, & Ormel, 2003; Krueger, Markon, Patrick, & Iacono, 2005). In Krueger et al. (2003), a two-factor internalizing/externalizing model best accounted for the data, with hazardous alcohol use being the only observed variable associated with the latent externalizing factor. Krueger et al. (2005) used latent class and latent trait methods to evaluate the underlying dimensionality associated with five disorders: conduct disorder (CD), adult antisocial behavior, alcohol dependence, cannabis dependence, and drug dependence. Their findings suggested that externalizing pathology is most accurately conceptualized as a continuous and normally distributed latent continuum of risk. Similar genetic modeling studies by Krueger et al. (2002) and Hicks, Krueger, Iacono, McGue, and Patrick (2004), based on symptom counts of four externalizing conditions (CD, adult antisocial behavior, alcohol dependence, and drug dependence) and the constraint personality domain, produced results consistent with a single underlying externalizing latent trait. Young, Stallings, Corley, Krauter, and Hewitt (2000) evaluated both phenotypic and genetic models of externalizing pathology and related personality features (CD, attention deficit/hyperactivity disorder [ADHD], substance experimentation, and novelty seeking), with findings from both analyses suggestive of a single latent factor.

The above findings on the modeling of externalizing symptoms and disorders contrast with other studies that have analyzed dimensionally-based personality or behavioral scales (Fergusson, Horwood, & Lynskey, 1994a, 1994b; Krueger, Markon, Patrick, Benning, & Kramer, 2007; Markon, Krueger & Watson, 2005) or symptom-level analyses associated with

childhood psychiatric disorders (e.g., Lahey et al., 2004, 2008). In Markon et al. (2005), for example, scores associated with 53 personality scales were analyzed. At the most superordinate level, two broad factors were identified, one roughly corresponding to neuroticism/disinhibition and the other to extraversion/openness. The neuroticism/disinhibition latent factor was further bifurcated into negative emotionality and disinhibition, with the disinhibition latent factor further divided into “disagreeable disinhibition” (the reverse of agreeableness) and “unconscientious disinhibition” (the reverse of conscientiousness) at lower levels in the hierarchy. Tackett, Krueger, Iacono, and McGue (2008) report a similar pattern of findings based on a middle childhood sample.

In Krueger et al. (2007), the hierarchical modeling of scores on 23 self-report scales related to externalizing tendencies was evaluated with exploratory and confirmatory analyses. Results indicated that a hierarchical two-factor model (substance use/theft/impulsivity and aggression/impatience/alienation) was superior to the general one-factor model. In Lahey et al. (2004), symptoms associated with CD were distinct from those that defined ADHD and oppositional defiant disorder (ODD), with symptom features of the latter two disorders demonstrating substantial overlap. Fergusson et al. (1994a) similarly reported that disruptive child behaviors were associated with two higher order factors defined by ODD/ADHD and CD-related behaviors.

Finally, in a large comprehensive study on the modeling of child and adolescent internalizing and externalizing symptoms, Lahey et al. (2008) identified a higher-order externalizing dimension that accounted for a significant proportion of variance in externalizing symptomatology (68 to 82%). Each increasingly differentiated nested model, however, fit the data significantly better than lesser-differentiated models. At intermediate levels in the hierarchy, the factor defined by CD remained distinct from those associated with ADHD (both inattention and impulsivity) and ODD. The best fitting model overall, however, consisted of four lower-order externalizing dimensions: inattention, hyperactivity, ODD symptoms, and CD symptoms. Although such dimensional studies provide evidence in support of lower-order factors linked to a higher-order externalizing latent factor, multifactor models of externalizing disorders that contain a substantial number of putative diagnostic categories have not been sufficiently evaluated, particularly with reference to patterns of lifetime diagnostic comorbidity.

This Study

The primary objective of the present research is to further evaluate and refine hierarchical measurement models of putative disorders from the externalizing domain. Specifically, this research is concerned with the evaluation and selection of theoretically plausible and competing measurement models of externalizing disorders based on patterns of lifetime comorbidity, and to evaluate the external validity of selected models when referenced to multiple indicators of adult psychosocial functioning.

The present study extends previous research in the following ways. First, this study incorporated a comparatively large number of *DSM*-defined putative externalizing disorders ($n = 7$): ADHD, ODD, CD, adult antisocial behavior (AAB; or the joint satisfaction of criterion A and B of *DSM-IV*-defined antisocial personality disorder), alcohol use disorder (ALC, defined by the presence of an alcohol abuse or dependence diagnosis), cannabis use disorder (CAN, as indicated by the presence of a cannabis abuse or dependence diagnosis), and hard drug use disorder (DRG, which includes the abuse or dependence of substances other than cannabis and alcohol).¹ Associations among nosological and statistically based paradigms are necessarily restricted by the range and breadth of diagnostic categories represented by each paradigm. Restricted statistical models that contain limited sets of putative disorders compared to those that appear in nosological systems can result in model fit indices that are optimal for

the limited set of disorders examined, but not reliable when additional disorder concepts that belong to the same family of disorders are added to the model (Wittchen, Hofler, & Merikangas, 1999). The inclusion of more putative externalizing disorders than found in previous studies allows for evaluations of measurement models that better represent the organizational structure proposed by different theoretical or nosological frameworks.

Second, in addition to testing a single-factor externalizing model, the present research evaluated several plausible multifactor models to account for observed patterns of lifetime disorder comorbidity. In so doing, the primary level of analysis in this research was on diagnostic categories and their associated patterns of lifetime comorbidity rather than covariation among symptom-based dimensions. Diagnostic criteria and disorder threshold cut-offs are related to clinical impressions about the severity of a disorder that, in turn, warrants treatment when the diagnosis is reached (Fergusson et al., 1994a). The present research, therefore, emphasizes the identification of distinct patterns of covariation that represent the extremes of broad underlying continuous variables.

Third, diagnostic membership was referenced to longitudinal data from an age-based cohort of community volunteers who were assessed on four separate occasions between the ages of 14 and 33. A potential confound in past similar research is that participants' ages usually varied considerably, even though some disorders are more prevalent during certain developmental periods (Wittchen et al., 1999). Additionally, in samples heterogeneous for age, the time period for lifetime disorder assessment is longer for older participants than for younger participants. Retrospective recalls of past events are known to be affected by a host of cognitive biases (Loftus & Pickrell, 1995), and there are indications that retrospective reports over extended time periods result in substantial underestimates of disorder occurrences (Masia et al., 2003; Wells & Horwood, 2004). The present research sought to minimize these potential sources of bias through the use of multiple diagnostic assessments on the same community-based cohort over a 19-year interval, a developmental period during which the first emergence of most psychiatric disorders is likely (Kessler et al., 1994; Kessler et al., 2005a).

Fourth, the validity of selected measurement models was further evaluated with respect to 17 psychosocial outcomes assessed at the last assessment wave. If more than one latent factor was identified among the resultant viable models, the external validity of the identified factors would be supported by discernable patterns of associated adult outcomes. The competing and theoretically plausible models of lifetime liability patterns of putative externalizing disorders evaluated in this research are pictorially summarized in Figure 1 and are described below.

EXT1—This model posits a single higher-order latent factor for the seven putative disorders (ADHD, ODD, CD, AAB, ALC, CAN, and DRG). This model is consistent with previous research on the genetic modeling of externalizing disorders (Hicks et al., 2004; Kendler et al., 2003; Krueger et al., 2002; Young et al., 2000) and studies where a single latent externalizing factor best accounted for the pattern of associations among a set of externalizing symptoms or disorders (Krueger, 1999; Krueger et al., 1998; Krueger et al., 2003; Slade & Watson, 2006; Vollebergh et al., 2001; see also meta-analysis reported in Krueger & Markon, 2006). The proposal of a single factor model, however, is in conflict with findings from other studies on

¹The use of the label “externalizing disorders” for the 7 disorder categories examined in the present research is consistent with descriptive terminology used in previous studies of these conditions (e.g., Krueger & Markon, 2006; Lahey et al., 2008). Externalizing behavior has historically been defined as a form of acting out behavior characterized by impulsive, hyperactive, aggressive, and rule-breaking acts. We note, however, that the traditional “externalizing” and “internalizing” distinction might have important limitations. For example, symptoms associated with major depressive disorder and generalized anxiety disorder, both typically associated with the internalizing domain, demonstrate substantial and essentially equal associations with symptoms of disorders associated with both internalizing and externalizing domains (Lahey et al., 2008). Additionally, although attention problems have historically been associated with the externalizing syndromes, some statistically-based frameworks have located difficulties in this area outside of this general domain (Achenbach, Bernstein, & Dumenci, 2005).

the dimensional modeling of symptoms and behaviors of externalizing disorders (reviewed below).

EXT2a—This two-factor model posits a latent factor that incorporates attention-deficit and disruptive behavior disorders (ADHD, ODD, CD, AAB) that is distinguished from another latent factor that accounts for substance use disorders (ALC, CAN, DRG). Except for the inclusion of AAB in the class of disorders defined by ADHD, ODD, and CD, this model is consistent with the structural organization of disorders in the *DSM-IV*, ADHD, ODD, and CD are regarded as belonging to the same general category of disorders (“attention-deficit and disruptive behavior disorders”), childhood CD is a pre-requisite for antisocial personality disorder (and, correspondingly, a risk factor for AAB), and substance use disorders are assigned to a category of disorders distinct from these other conditions.

EXT2b—This posited two-factor model distinguishes oppositional behavior disorders (ADHD, ODD) from social norm violation disorders (CD, AAB, ALC, CAN, DRG). This model is influenced by studies that report considerable covariation of ADHD and ODD symptoms and latent factors (Burns, Boe, Walsh, Sommers-Flanagan & Teegarden, 2001; Burns, Walsh, Owen, & Snell, 1997a; Burns et al., 1997b; Lahey et al., 2004; Lahey et al., 2008), extensive overlap in the genetic and environmental influences associated with ADHD and ODD (Burt, Krueger, McGue, & Iacono, 2003), clear distinctions among ODD and ADHD symptoms from CD symptoms (Burns et al., 1997a; Burns et al., 1997b; Fergusson et al., 1994a; Frick et al., 1993; Lahey et al., 2004; Lahey et al., 2008), as well as reports of significant associations of CD and/or AAB symptoms with substance abuse based on phenotypic (Fergusson, Horwood, & Lynskey, 1993; Frick et al., 1993) and genetic (Hicks et al., 2004; Kendler et al., 2003; Krueger et al., 2002) features. This modeling is also consistent with research that suggests CD, antisocial personality disorder, ALC, CAN and/or DRG diagnoses are best accounted for by a single latent factor (e.g., Krueger, 1999; Krueger et al., 1998; Krueger et al., 2003; Slade & Watson, 2006; Vollebergh et al., 2001).

EXT3—This model posits three latent externalizing factors: oppositional behavior disorders (ADHD, ODD), social norm violation disorders (CD, AAB), and substance use related disorders (ALC, CAN, DRG). This model is consistent with research reviewed in relation to the EXT2b model. The main difference, however, is that substance use disorders are distinguished from CD and AAB and defined by a separate latent factor. This model is also in line with findings that suggest alcohol and drug abuse/dependence are associated with various forms of disinhibitory personality characteristics in addition to CD and AAB (Grant et al., 2004; Sher & Trull, 1994), very high rates of co-occurring abuse across drug categories (Tsuang et al., 1998), and research by Kendler et al. (2003) who identified specific genetic and environmental risk factors for drug and alcohol dependence that were distinct from those associated with CD and AAB.

METHOD

Participants

Participants in this research consisted of 816 individuals (59% women, 41% men; 89% White) who began participation in the Oregon Adolescent Depression Project during mid-adolescence and continued with the study through age 33. Detailed descriptions of this sample have been provided elsewhere (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993; Lewinsohn, Rohde, Seeley, Klein, & Gotlib, 2003; Rohde et al., 2007). Briefly, 1,709 adolescents between the ages of 14 and 19 were randomly selected from 9 high schools that were representative of urban and rural districts in western Oregon. Participants were initially assessed between 1987 and 1989 (T_1), and 1,507 of these persons were again assessed about a year ($M = 13.8$ months,

$SD = 2.3$) later (T_2). The mean ages of youth at T_1 and T_2 were 16.6 ($SD = 1.2$) and 17.7 ($SD = 1.2$) years, respectively.

A third assessment wave (T_3) was conducted between 1994 and 1999 as participants reached their 24th birthday. Eligible participants for this third assessment wave included all persons with a positive history of a psychiatric diagnosis at T_2 ($N = 644$) and a randomly selected subset of participants with no history of mental disorder at T_2 ($N = 457$ of 863 persons, which also included all non-White T_2 participants in order to enrich the ethnic representativeness and diversity of the sample). Of these 1,101 eligible persons, 941 (85%) completed the T_3 diagnostic interview an average of 6.8 years ($SD = 1.4$) following the T_2 assessment.

A fourth assessment wave (T_4) was conducted between 2000 and 2004 as participants reached their 30th birthday. From the 941 eligible persons who completed T_3 , 816 (87%) completed the T_4 assessment. Of these, 348 (42.6%) were among the 457 randomly selected persons at T_3 from the pool of persons not diagnosed with a lifetime psychiatric disorder by T_2 . The remaining 468 persons (57.4%) were among the 644 T_2 participants with a positive history of at least one psychiatric diagnosis by T_2 . The mean age of all T_4 participants was 30.6 ($SD = 0.6$), and 41% had earned a bachelor's degree or higher. Almost half of the sample (53%) was married at the time of assessment. These 816 persons constitute the reference sample for this research.

Rohde et al. (2007) review findings related to participant attrition over the four assessment waves. Overall, attrition between T_1 and T_4 was modest. Between T_1 and T_2 , some statistically significant differences between those who continued and discontinued participation were observed, although these differences were relatively small in magnitude. Discontinuation of participation was more common among males (6.4%) than females (5.4%), and associated with lower socioeconomic status, fewer people living at home, current and past cigarette use, a prior substance use disorder, and a past diagnosis of a disruptive behavior disorder. At T_3 , men were more likely to discontinue participation (19%) than women (11%). Diagnostic status or demographic variables assessed at T_2 were not significantly related to attrition by T_3 . By T_4 , discontinuation was more common among men (16%) than women (11%), and among individuals with a lifetime diagnosis of a substance abuse disorder by T_3 .

Diagnostic Interviews

During the first three assessment waves, participants were interviewed with a version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS) that combined features of the Epidemiologic and Present Episode versions (Chambers et al., 1985; Orvaschel, Puig-Antich, Chambers, Tabrizi, & Johnson, 1982). Follow-up assessments of disorders at T_2 and T_3 also involved the joint administration of the Longitudinal Interval Follow-Up Evaluation (LIFE; Keller et al., 1987) that, in conjunction with the K-SADS, provided detailed information related to the presence and course of disorders since participation in the previous diagnostic interview. The T_4 assessment included administration of the LIFE and the Structured Clinical Interview for Axis I *DSM-IV* Disorders–Non-Patient Edition (SCID-NP; First, Spitzer, Gibbon, & Williams, 1994). AAB was assessed with the Personality Disorder Examination (Loranger, 1988) at T_3 and the International Personality Disorder Examination (IPDE; Loranger et al., 1994) at T_4 . Diagnostic categories were evaluated in accordance with *DSM-III-R* criteria (American Psychiatric Association, 1987) at T_1 and T_2 and *DSM-IV* criteria (American Psychiatric Association, 1994) at T_3 and T_4 .

Following a methodology utilized in previous studies (Krueger, 1999; Krueger et al., 1998; Vollebergh et al., 2001), disorders that had very low lifetime incidence ($< 1\%$) were excluded from further consideration. Disorders included in the present research (with associated lifetime

prevalence rates through T₄) were: ADHD (3.2%), ODD (3.4%), CD (5.0%), AAB (5.3%), ALC (37.4%), CAN (21.2%), and DRG (13.8%).

ADHD, CD, ALC, CAN, and DRG were assessed at each of the four waves. ODD was assessed from T₁ through T₃; however, neither ODD nor CD diagnoses were made at T₃ or T₄. In instances where symptom criteria for these disorders were met, a diagnosis of antisocial personality disorder (APD) was determined to be more appropriate according to *DSM-IV* decision rules. Because the *DSM-IV* criteria for APD include a childhood history of CD, thus creating a spurious dependency between these two diagnostic categories, only criterion A (adult antisocial behavior, or AAB) and criterion B (age $\geq$ 18) of APD were considered in relation to the assessment of AAB in this research. Those who jointly satisfied both of these criteria at either T₃ or T₄ received a diagnosis of AAB.

Interviewers for each of the four assessment waves were carefully trained and supervised, with most having advanced degrees in mental health-related disciplines (see Rohde et al., 2007, for details on assessor training and supervision). All interviews were either audio- or video-taped, and interviews from each assessment wave were randomly selected for reliability assessments by a second interviewer.

Inter-rater reliability for diagnostic categories was indexed by kappa. In order to avoid the potential inflation, deflation, and/or unreliability of the kappa statistic that can occur when few positive cases are observed, kappa coefficients were only computed for categories diagnosed as present at least 10 times when summed over the two raters. Diagnostic reliability was good to excellent across the four assessment waves. Two disorders, ALC and CAN, were diagnosed with sufficient frequency among raters during 3 of the 4 waves, and the mean kappas were, respectively, .77 (range: .74 to .82) and .79 (range: .72 to .83). AAB and DRG were diagnosed with sufficient frequency during 2 of the 4 waves, and the mean kappa for these disorders were .61 (range: .49 to .74) and .76 (range: .69 to .83), respectively. Kappa coefficients for ADHD (.89) and ODD (.77) could only be determined for 1 of the 4 assessment waves. CD was not diagnosed with sufficient frequency during any of the assessment waves to allow an evaluation of inter-rater diagnostic reliability.²

Psychosocial Functioning at T₄ and Lifetime History of Mood and Anxiety Disorders

Viable measurement models of putative externalizing disorders, as determined by procedures described in the following section, were further evaluated in terms of their ability to account for variance in several indicators of psychosocial functioning assessed at T₄ as well as lifetime histories of depressive and anxiety disorders. At T₄, participants completed a survey that included assessments of demographic information (i.e., years of education, current employment status, and annual household income) and psychosocial functioning. These latter areas included marital history (i.e., never married, married but subsequently divorced/separated, currently married and never divorced/separated), history of biological parentage of a child, relationship quality with family and friends, and level of social adjustment during the past two weeks. General health functioning (i.e., self-rated physical health, engagement in high-risk sexual behavior during the past year) and broad indicators of psychological functioning (i.e., suicide attempt between T₃ and T₄ assessments, self-ratings of life dissatisfaction, coping skills and resources, and stressful major life events in the last year) were also surveyed. Detailed descriptions of these variables can be found in Rohde et al. (2007).

Based on data from diagnostic interviews collected from T₁ through T₄, we also evaluated the

²Although it was not possible to compute kappa coefficients for CD, symptom ratings for disorder criteria were available for two independent raters from the T₁ assessment. An intraclass correlation coefficient (ICC) was computed to reference rater agreement based on symptom counts, and the corresponding value was .91. This ICC value exceeds the threshold suggested by Fleiss (1981) for excellent agreement (.75).

associations that latent factors from selected models had with lifetime histories of depressive disorders (i.e., major depressive disorder, dysthymia) and anxiety disorders (i.e., separation anxiety, generalized anxiety, social phobia, single/specific phobias, agoraphobia, panic, obsessive-compulsive, acute stress, and post-traumatic stress).

Statistical Analyses and Model Selection

There is growing consensus that goodness-of-fit is a necessary but not sufficient condition in model selection, and that the issue of model parsimony versus complexity (and hence generalizability) must also be considered (Myung, 2000). Accordingly, model evaluation and selection in the present research were jointly influenced by both of these considerations.

Evaluation of goodness-of-fit—CFA was used to evaluate the goodness-of-fit of competing externalizing models, with model parameters estimated with the weighted least squares estimator with means and variances adjusted (WLSMV) procedure. This estimation method was selected for two reasons. First, maximum likelihood (ML) estimation procedures that result in a standard log likelihood ratio chi-square test of model fit are not robust when data are sparse (see Fergusson & Horwood, 2002). Second, unlike the ML approach, the WLSMV method does not assume joint multivariate normality of disorder distributions within the population sampled, and is robust against violations of distribution normality in samples of varying size (Curran, West, & Finch, 1996; Flora & Curran, 2004). Mplus statistical software (version 5.1; Muthén & Muthén, 1998-2007) was used to perform all CFA analyses.

In the evaluation of model fit, several indices were used and interpreted in conjunction with critical cut points recommended by Yu (2002). These fit indicators (with cut-off values in parentheses) were: chi-square test of model fit (X^2 ; $p > .05$), comparative fit index (CFI; $\geq .96$), Tucker–Lewis fit index (TLI; $\geq .95$), root mean square error of approximation (RMSEA; $\leq .05$), and weighted root mean square residual (WRMR; ≤ 1.0).

Evaluation of model complexity—Comparisons were also made to identify which theoretically plausible models best approximated underlying processes contained in the data. The information-theoretic approach to model selection (e.g., Burnham & Anderson, 2004; Myung, 2000) employs the use of information criterion statistics that reference the optimal balance of model fit with parsimony. Several information-theoretic statistics for model selection are available to researchers, including the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) statistics. These statistics primarily differ in terms of how complexity is measured and combined with goodness-of-fit indicators in the computation of an overall selection criterion (Myung, 2000). BIC compared to AIC favors more parsimonious models when sample sizes are small to moderate by penalizing additional estimated model parameters and is, consequently, more likely to result in the selection of an underfitted model (Burnham & Anderson, 2004). Underfitted models, some have argued, are a more serious problem than over-fitted models in data analysis and inference (e.g., Shibata, 1989). Consequently, the sample-size adjusted BIC, which does not impose as harsh a penalty on additional parameters and is a more accurate statistic than BIC under some circumstances (Henson, Reise, & Kim, 2007; Yang, 2006), was used in addition to AIC to inform model selection.

In the interpretation of AIC and adjusted BIC statistics, lower values indicate greater parsimony and fit, or an optimal balance between under- and over-fitted models. There are currently no inferential probability statements, such as a p -value, available to evaluate confidence in magnitude differences in these criteria. Burnham and Anderson (2004), however, offer general recommendations for interpreting AIC values. Difference scores (denoted as ΔAIC) were derived by subtracting the AIC value associated with the best approximating model (i.e., the

one with the lowest AIC value) within a set from AIC values of other models within the same set. Following Burnham and Anderson's guidelines, ΔAIC values of ≤ 2.0 are indicative of similar competing models in terms of their approximating abilities. ΔAIC values between 4.0 and 7.0 indicate models that have considerably less support, and models with values > 10.0 have essentially no support relative to the best approximating model. Raftery (1995) has offered similar guidelines for non-adjusted BIC; however, there are presently no comparable guidelines for adjusted BIC. Model selection in the present study, therefore, was guided by (a) fit statistics from CFA analyses, (b) AIC and adjusted BIC values, and (c) interpretative guidelines provided by Burnham and Anderson based on ΔAIC values.

Caucasian T_2 participants without a history of psychiatric diagnosis were under-sampled in subsequent assessment waves due to stratified sampling procedures implemented at T_3 . To adjust for this, a weighting procedure was used in some analyses whereby Caucasian participants without a lifetime diagnosis by T_2 were assigned a weight of 2.05, a value that reflects the probability of this subgroup being sampled at T_3 . In contrast, all participants with a psychiatric diagnosis by T_2 and all non-Caucasian participants were assigned a weighting of 1.0. In the following sections, analyses based weighted data are presented.

RESULTS

Tetrachoric Correlations and Relative Risk Comparisons among Diagnostic Categories

Table 1 presents tetrachoric correlations as approximations of the relative lifetime risk among diagnostic categories examined in this research. Correlations among the externalizing disorders were consistently positive and frequently moderate to large in magnitude. Particularly noteworthy are the comparatively modest associations that ADHD and ODD have with the various substance use disorders, and the relatively large associations shared among the substance use disorders. CD and AAB associations with substance use disorder were consistently moderate in magnitude (*range*: .47 to .64). Overall, when comparable disorders were evaluated, correlations reported here are similar in magnitude to those presented in related studies (Krueger, 1999; Slade & Watson, 2006).

Odds ratios also displayed in Table 1 generally suggest that the presence of any externalizing disorder is associated with a substantial increased risk for other disorders within this domain. When the relative risk among pairs of specific disorders is considered, however, 95% confidence intervals in some cases include values that are ≤ 1.0 . In such instances, the presence of a given disorder cannot be regarded as significantly adding to the relative risk of another disorder. Overall, ADHD and ODD tended to have smaller odds ratios and a greater number of confidence intervals that included values ≤ 1.0 compared to other disorders examined.

Evaluation of Externalizing Models

Table 2 presents fit indicators and selection criteria for each of the models evaluated. Fit indicators were generally satisfactory across models. The exception was the X^2 test, which suggested an unsatisfactory model fit for EXT1. All fit indicators for EXT2a, EXT2b, and EXT3 suggested good model fits.

Information-theoretic criteria presented in Table 2 indicate that, of the four competing models evaluated, EXT2b fit the data best ($\text{AIC} = 44.98$, adjusted $\text{BIC} = -20.31$). When referenced to the interpretative guidelines provided by Burnham and Anderson (2004), ΔAIC criteria suggest that none of the remaining three models had approximating abilities similar to EXT2b.

Figure 2 presents a path diagram of EXT2b. In this figure, parameter estimates are standardized. As depicted in Figure 2, CD, AAB, ALC, CAN, and DRG are all highly related to the same latent factor (labeled "social norm violation disorders"). ADHD and ODD, in contrast, were

modeled as belonging to a distinct latent factor (labeled “oppositional behavior disorders”), with both disorders demonstrating either moderate or large associations with this factor. The two latent factors also demonstrated a moderate association with each other ($r = .46, p = .002$). In the aggregate, these findings suggest that although these 7 putative externalizing disorders are best modeled as two distinct factors, there are nonetheless some shared features among these factors.

Associations with Psychosocial Functioning Indicators at T₄

In a series of analyses, each outcome measure of psychosocial functioning assessed at T₄ was regressed on the latent factors associated with EXT2b. Standardized regression coefficients are presented in Table 3. With respect to model EXT2b, several outcome variables demonstrated statistically significant and distinct relationships with the two latent factors. Specifically, the oppositional behavior disorders factor is uniquely and negatively associated with years of schooling. In contrast, the social norm violation disorders factor is uniquely associated with several indicators of impaired functioning. These indicators included comparatively low income, current non-married status with a greater history of never being married, impaired social adjustment, poorer physical health, a greater tendency to engage in high risk sexual behavior, a greater likelihood of at least one suicide attempt since T₃, more life dissatisfaction, comparatively poor coping skills and resources, greater exposure to stressful life events, and increased lifetime risk for depressive disorders.

DISCUSSION

A primary goal of the present study was to evaluate several theoretically plausible and competing models of lifetime comorbidity among the externalizing domain of disorders. In contrast to previous studies in which a single externalizing latent factor best accounted for these disorders (Krueger, 1999; Krueger et al., 1998; Krueger et al., 2003; Krueger & Markon, 2006; Slade & Watson, 2006; Vollebergh et al., 2001), the present research suggests that the seven disorders examined here are best modeled with two latent factors. The discrepant findings between this research and similar prior studies are likely related to several factors. In prior studies of diagnostic categories, for example, four or fewer externalizing disorder candidates were modeled. The majority of these studies also evaluated disorder presence within a narrow time frame, usually within a year of the diagnostic assessment, thus reducing the likelihood that alternative manifestations of the underlying liability would be expressed or detected.

Each of the four models evaluated in this study generally demonstrated satisfactory model fit on a majority of standard fit indicators. EXT1, however, just failed to reach the conventional threshold for good fit on the X^2 goodness-of-fit test, which is known to produce significant p -values even when the underlying model is reasonable (Bentler, 1990). Apart from this stringent test, other indicators suggested that this model along with EXT2a and EXT3 produced reasonable fits to the data. These models, however, were not as effective in accounting for the data as EXT2b according to adjusted BIC, AIC, and Δ AIC criteria.

The best fitting two-factor model (EXT2b) distinguished ADHD/ODD from CD/AAB/ALC/CAN/DRG. This two-factor solution partially replicates higher-order factor models obtained by Lahey et al. (2004, 2008), in which symptoms and behaviors associated with ADHD and ODD loaded on the same factor and were distinct from those that defined CD, which loaded on a separate correlated factor. Similarly, this solution also partially replicates and extends findings from other modeling studies that suggest CD, AAB, ALC, CAN and DRG diagnoses are best accounted for by a single latent factor (Krueger, 1999; Krueger et al., 1998; Krueger et al., 2003; Slade & Watson, 2006; Vollebergh et al., 2001).

The correlated latent factors of the selected two-factor model also show some correspondence with the inattentive/hyperactive/impulsive versus conduct problems/rule-breaking/violent behavior distinctions that are emerging in the research literature (e.g., Waschbusch, 2002). The latent factor associated with ADHD/ODD in EXT2b is likely characterized by inattention and tendencies to display intrusive, impulsive, and hyperactive behaviors (Lahey et al., 2004). The comparatively modest correlations and odds ratios that ADHD and ODD had with substance abuse disorders relative to CD and AAB are consistent with findings from other epidemiological studies in which conduct problems were found to be more strongly associated with future substance abuse problems than attention deficit-related problems (Lynskey & Fergusson, 1995).

The second latent factor defined by CD/AAB/ALC/CAN/DRG likely reflects a general rule-breaking tendency. CD, AAB, and substance use disorders share similar risk factors, such as an impulsive temperament, affiliations with rule-breaking or substance using peers, deviant family systems, and parental use of substances (Burt et al., 2001; Fergusson & Horwood, 1999; Lynskey, Fergusson, & Horwood, 1998; Wills & Dishion, 2004). Early conduct problems alone appear to explain much of the future risk for substance abuse (Lynskey & Fergusson, 1995), an observation that may explain their emergence on the same latent factor in the present research. Similarly, the observation that each of the substance use categories (ALC, CAN, and DRG) define, in part, this latent factor is also consistent with research that demonstrates a high degree of comorbidity among substance abuse disorders (Tsuang et al., 1998).

Support for the validity of the two latent factors in model EXT2b is evident in comparisons of psychosocial outcomes at T₄. When compared to the social norm violation latent factor, the oppositional behavior disorders latent factor was associated with comparatively better overall psychosocial functioning. The main distinguishing feature of this latent factor was its association with fewer years of schooling. This finding is consistent with other research that associates learning problems and impaired academic performance with inattentiveness and hyperactive and oppositional behavioral tendencies (Fischer, Barkley, Edelbrock, & Smallish, 1990). In contrast, the social norm violation latent factor was associated with several negative outcomes, findings that are consistent with the observed cumulative effects of antisocial developmental pathways (Moffitt, Caspi, Harrington, & Milne, 2002).

Although findings from this research are consistent with those from other modeling studies (e.g., Burns et al., 1997a; Burns et al., 1997b; Fergusson et al., 1994a; Lahey et al., 2004; Lahey et al., 2008), they depart somewhat from the organization of disorders in the *DSM-IV* framework. In *DSM*, ADHD, ODD, and CD are listed as diagnostic categories within the same section ("attention-deficit and disruptive behavior disorders"). Although the present research and that performed by others suggest that the lifetime risk of these three disorders are greater than would be suggested by chance alone, ADHD and ODD have been shown to be distinct from CD in several respects. ADHD and ODD, for example, demonstrate substantial symptom covariation, diagnostic comorbidity, and common genetic and environmental determinants (Burns et al., 2001; Burns et al., 1997a; Burns et al., 1997b; Burt et al., 2003; Lahey et al., 2004; Lahey et al., 2008), with these two disorders clearly distinguished from CD in these areas (e.g., Burns et al., 1997a; Fergusson et al., 1994a; Lahey et al., 2008). Also in conflict with the *DSM* framework, substance use disorders were not found to be distinct from the disruptive behavior problems that define CD and AAB in the present research.

Overall, findings from this study support the proposition of a multifactor externalizing domain of psychological disorders. Persons elevated on this general liability are at greater risk for expressing or manifesting multiple disorders within this domain, particularly disorders that share common latent factors at lower levels in the hierarchy. These latent factors, in turn, might reflect the overall effects of genetic, physiological, and environmental influences on individual

vulnerabilities to externalizing symptoms and behaviors (Fergusson, Horwood, & Boden, 2006). Included among these influences is neurochemical activity. Serotonergic functions, for example, are inversely related to externalizing behavior tendencies while dopaminergic functions are positively associated with such tendencies (Chambers, Taylor, & Potenza, 2003). Other etiological factors might include reduced cognitive ability (Koenen, Caspi, Moffitt, Rijdsdijk, & Taylor, 2006; Molina, Smith, & Pelham, 2001), impairments in the ability to inhibit or override prepotent responses resulting in behavioral inflexibility and perseveration (Nigg, 2001), heightened sensitivity and responsiveness to reward cues (Iacono, Malone, & McGue, 2008; Nigg, 2001), and difficulties linking consequences to the behaviors that produced them (Iacono et al., 2008). Other potential risk factors for disruptive behavior and substance use disorders include a childhood history of neglect or abuse, poor or inconsistent parenting, and affiliations with deviant peers (Appleyard, Egeland, van Dulmen, & Sroufe, 2005; Iacono et al., 2008).

The identification of latent factors associated with the emergence of externalizing disorders during adolescent and adult development is likely to facilitate additional research on central pathological processes (Krueger, 1999) or common core diagnostic features (Wittchen et al., 1999). To the extent that there are genetic underpinnings associated with psychiatric diagnoses, for example, these are more likely to be associated with latent factors than specific diagnostic categories (Hettema, Neale, Myers, Prescott, & Kendler, 2006).

Study Limitations

Findings and conclusions associated with this research should be considered along with some caveats. First, as with many longitudinal studies, several modifications in procedures were made over the study period that might have impacted research findings. For example, *DSM* disorder criteria and diagnostic decision rules changed from T_1/T_2 (*DSM-III-R*) to T_3/T_4 (*DSM-IV*). Similarly, the diagnostic interview used in assessments from T_1 to T_3 (K-SADS, PDE) differed from that used at T_4 (SCID, IPDE). Variability in diagnostic criteria and interview formats across assessment waves may have introduced some method bias that, in turn, altered associations among some subsets of psychiatric disorders.

Second, research participants were ethnically and geographically homogeneous. Previous research in this area has pointed to some cultural and geographical differences in the hierarchical organization of the internalizing spectrum of psychiatric disorders (Krueger et al., 2003). Consequently, the generalizability of findings obtained in this research to members of diverse ethnic or cultural groups or to persons from diverse geographic regions remains uncertain.

Third, although this research spanned a developmentally important age range in relation to the emergence of psychological disorders, patterns of associations among some diagnostic categories might change with age. Even though the vast majority of cases of alcoholism emerge before age 30 (Helzer et al., 1990), for example, there are indications that persons who first develop this condition later in life demonstrate fewer externalizing tendencies compared to those who initially develop this condition at younger ages (Windle & Scheidt, 2004). Accordingly, the extent to which the measurement models presented here generalize to middle and late adulthood is unclear.

Future Directions

Clark (2005) has suggested that a comprehensive model of psychopathology should delineate the relationships among disorder concepts, account for non-random comorbidity among disorders, illustrate the temporal sequencing of broadband or superordinate personality traits or liabilities (e.g., temperament) with the emergence of behavioral patterns or psychiatric

disorders, and demonstrate consistencies with emerging research on developmental factors, both biological and environmental, that are causally related to disorders and broad personality domains subsumed by the model. The present study attempted to directly address the first three of these research goals, although we acknowledge there is room for the further development, refinement, and evaluation of the models presented here. For example, other candidate disorders not included in this study might be pertinent for an externalizing spectrum of disorders (e.g., narcissistic, borderline, and histrionic personality disorders; impulse control disorders). Additional tests of these models are needed, including those that involve genetic and environmental risk modeling, dimensional representation of symptom features, and evaluations of differential responsiveness to therapies for closely aligned disorders.

Acknowledgments

National Institute on Drug Abuse Grant DA12951 and National Institute of Mental Health Grants MH40501 and MH50522 supported this research.

References

- Achenbach TM, Bernstein A, Dumenci L. *DSM*-oriented scales and statistically based syndromes for ages 18 to 59: Linking taxonomic paradigms to facilitate multitaxonomic approaches. *Journal of Personality Assessment* 2005;84:49–63. [PubMed: 15639767]
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 3. Washington, D.C.: Author; 1987.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4. Washington, DC: Author; 1994.
- Appleyard K, Egeland B, van Dulmen MHM, Sroufe LA. When more is not better: The role of cumulative risk in child behavior outcomes. *Journal of Child Psychology and Psychiatry* 2005;46:235–245. [PubMed: 15755300]
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin* 1990;107:238–246. [PubMed: 2320703]
- Burnham KP, Anderson DR. Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research* 2004;33:261–304.
- Burns GL, Boe B, Walsh JA, Sommers-Flanagan R, Teegarden LA. A confirmatory factor analysis on the *DSM-IV* ADHD and ODD symptoms: What is the best model for the organization of these symptoms? *Journal of Abnormal Child Psychology* 2001;29:339–349. [PubMed: 11523839]
- Burns GL, Walsh JA, Owen SM, Snell J. Internal validity of attention deficit hyperactivity disorder, oppositional defiant disorder, and overt conduct disorder symptoms in young children: Implications from teacher ratings for a dimensional approach to symptom validity. *Journal of Clinical Child Psychology* 1997a;26:266–275. [PubMed: 9292384]
- Burns GL, Walsh JA, Patterson DR, Holte CS, Sommers-Flanagan R, Parker CM. Internal validity of the disruptive behavior disorder symptoms: Implications from parent ratings for a dimensional approach to symptom validity. *Journal of Abnormal Child Psychology* 1997b;25:307–319. [PubMed: 9304447]
- Burt SA, Krueger RF, McGue M, Iacono W. Parent-child conflict and the comorbidity among childhood externalizing disorders. *Archives of General Psychiatry* 2003;60:505–513. [PubMed: 12742872]
- Chambers RA, Taylor JR, Potenza MN. Developmental neurocircuitry of motivation in adolescence: A critical period of addition vulnerability. *American Journal of Psychiatry* 2003;160:1041–1052. [PubMed: 12777258]
- Chambers WJ, Puig-Antich J, Hirsch M, Paez P, Ambrosini PJ, Tabrizi MA, et al. The assessment of affective disorders in children and adolescents by semistructured interview: Test-retest reliability of the Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present Episode version. *Archives of General Psychiatry* 1985;42:696–702. [PubMed: 4015311]
- Clark LA. Temperament as a unifying basis for personality and psychopathology. *Journal of Abnormal Psychology* 2005;114:505–521. [PubMed: 16351374]

- Cox BJ, Clara IP, Enns MW. Posttraumatic stress disorder and the structure of common mental disorders. *Depression and Anxiety* 2002;15:168–171. [PubMed: 12112721]
- Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods* 1996;1:16–29.
- Fergusson DM, Horwood LJ. Prospective childhood predictors of deviant peer affiliations in adolescence. *Journal of Child Psychology and Psychiatry* 1998;40:581–592. [PubMed: 10357164]
- Fergusson DM, Horwood LJ. Male and female offending trajectories. *Development and Psychopathology* 2002;14:159–177. [PubMed: 11893091]
- Fergusson DM, Horwood LJ, Boden JM. Structure of internalising symptoms in early adulthood. *British Journal of Psychiatry* 2006;189:540–546. [PubMed: 17139039]
- Fergusson DM, Horwood LJ, Lynskey MT. Prevalence and comorbidity of DSM-III-R diagnoses in a birth cohort of 15 year olds. *Journal of the American Academy of Child and Adolescent Psychiatry* 1993;32:1127–1134. [PubMed: 8282656]
- Fergusson DM, Horwood LJ, Lynskey MT. Structure of DSM-III-R criteria for disruptive behaviors: Confirmatory factor models. *Journal of the American Academy of Child and Adolescent Psychiatry* 1994a;33:1145–1155. [PubMed: 7982865]
- Fergusson DM, Horwood LJ, Lynskey MT. The comorbidities of adolescent problem behaviors: A latent class model. *Journal of Abnormal Child Psychology* 1994b;22:339–354. [PubMed: 8064037]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for Axis I DSM-IV Disorders–Non-Patient Edition. New York: Biometrics Research Department; 1994.
- Fischer M, Barkley RA, Edelbrock CS, Smallish L. The adolescent outcome of hyperactive children diagnosed by research criteria: II. Academic, attentional, and neuropsychological status. *Journal of Consulting and Clinical Psychology* 1990;58:580–588. [PubMed: 2254504]
- Fleiss, JL. Statistical methods for rates and proportions. 2. New York: Wiley; 1981.
- Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 2004;9:466–491. [PubMed: 15598100]
- Frick PJ, Van Horn Y, Lahey BB, Christ MAG, Loeber R, Hart EA, et al. Oppositional defiant disorder and conduct disorder: A meta-analytic review of factor analyses and cross-validation in clinic sample. *Clinical Psychology Review* 1993;13:319–340.
- Grant BF, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry* 2004;61:361–368. [PubMed: 15066894]
- Helzer JE, Canino GJ, Yeh E-K, Bland RC, Lee CK, Hwu H-G, Newman S. Alcoholism–North American and Asia: A comparison of population surveys with the Diagnostic Interview Schedule. *Archives of General Psychiatry* 1990;47:313–319. [PubMed: 2322082]
- Henson JM, Reise SP, Kim KH. Detecting mixtures from structural differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling* 2007;14:202–226.
- Hettema JM, Neale MC, Myers JM, Prescott CA, Kendler KS. A population-based twin study of the relationship between neuroticism and internalizing disorders. *American Journal of Psychiatry* 2006;163:857–864. [PubMed: 16648327]
- Hicks BM, Krueger RF, Iacono WG, McGue M, Patrick CJ. Family transmission and heritability of externalizing disorders: A twin-family study. *Archives of General Psychiatry* 2004;61:922–928. [PubMed: 15351771]
- Iacono WG, Malone SM, McGue M. Behavioral disinhibition and the development of early-onset addiction: Common and specific influences. *Annual Review of Clinical Psychology* 2008;4:325–348.
- Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott PA. The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* 1987;44:540–548. [PubMed: 3579500]
- Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry* 2003;60:929–937. [PubMed: 12963675]

- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of *DSM-IV* disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 2005a;62:593–602. [PubMed: 15939837]
- Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of 12-month *DSM-IV* disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 2005b; 62:617–627. [PubMed: 15939839]
- Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry* 1994;51:8–19. [PubMed: 8279933]
- Koenen KC, Caspi A, Moffitt TE, Rijdsdijk F, Taylor A. Genetic influences on the overlap between low IQ and antisocial behavior in young children. *Journal of Abnormal Psychology* 2006;115:787–797. [PubMed: 17100536]
- Krueger RF. The structure of common mental disorders. *Archives of General Psychiatry* 1999;56:921–926. [PubMed: 10530634]
- Krueger RF, Caspi A, Moffitt TE, Silva PA. The structure and stability of common mental disorders (*DSM-III-R*): A longitudinal–epidemiological study. *Journal of Abnormal Psychology* 1998;107:216–227. [PubMed: 9604551]
- Krueger RF, Chentsova-Dutton YE, Markon KE, Goldberg D, Ormel J. A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *Journal of Abnormal Psychology* 2003;112:437–447. [PubMed: 12943022]
- 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. Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology* 2006;2:111–133.
- 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]
- Lahey BB, Applegate B, Waldman ID, Loft JD, Hankin BL, Rick J. The structure of child and adolescent psychopathology: Generating new hypotheses. *Journal of Abnormal Psychology* 2004;113:358–385. [PubMed: 15311983]
- Lahey BB, Rathouz PJ, Van Hulle C, Urbano RC, Krueger RF, Applegate B, 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]
- Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. Prevalence and incidence of depression and other *DSM-III-R* disorders in high school students. *Journal of Abnormal Psychology* 1993;102:133–144. [PubMed: 8436689]
- Lewinsohn PM, Rohde P, Seeley JR, Klein DN, Gotlib IH. Psychosocial functioning of young adults who have experienced recovery from major depressive disorder during adolescence. *Journal of Abnormal Psychology* 2003;112:353–363. [PubMed: 12943014]
- Loftus EF, Pickrell JE. The formation of false memories. *Psychiatric Annals* 1995;25:720–725.
- Loranger, AW. *Personality Disorder Examination (PDE) manual*. Yonkers, NY: DV Communications; 1988.
- Loranger AW, Sartorius N, Andreoli A, Berger P, Buchheim P, Channabasavanna SM, et al. The International Personality Disorder Examination: The World Health Organization/Alcohol, Drug Abuse, and Mental Health Administration international study of personality disorders. *Archives of General Psychiatry* 1994;51:215–224. [PubMed: 8122958]
- Lynskey MT, Fergusson DM. Childhood conduct problems, attention deficit behaviors, and adolescent alcohol, tobacco, and illicit drug use. *Journal of Abnormal Child Psychology* 1998;23:281–302. [PubMed: 7642838]

- Lynskey MT, Fergusson DM, Horwood LJ. The origins of the correlations between tobacco, alcohol, and cannabis use during adolescence. *Journal of Child Psychology and Psychiatry* 1998;39:995–1005. [PubMed: 9804032]
- Markon KE, Krueger RF, Watson D. Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology* 2005;88:139–157. [PubMed: 15631580]
- Masia CL, Storch EA, Dent HC, Adams P, Verdelli H, Davies M, et al. Recall of childhood psychopathology more than 10 years later. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003;42:6–12. [PubMed: 12500071]
- Moffitt TE, Caspi A, Harrington H, Milne BJ. Males on the life-course-persistent and adolescent-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology* 2002;14:179–207. [PubMed: 11893092]
- Molina BSG, Smith BH, Pelham WE. Factor structure and criterion validity of secondary school teacher ratings of ADHD and ODD. *Journal of Abnormal Child Psychology* 2001;29:71–82. [PubMed: 11316336]
- Muthén, LK.; Muthén, BO. *Mplus User's Guide*. 5. Los Angeles, CA: Muthén and Muthén; 1998/2007.
- Myung IJ. The importance of complexity in model selection. *Journal of Mathematical Psychology* 2000;44:190–204. [PubMed: 10733864]
- Nigg J. Is ADHD a disinhibitory disorder? *Psychological Bulletin* 2001;127:571–598. [PubMed: 11548968]
- Orvaschel H, Puig-Antich J, Chambers WJ, Tabrizi MA, Johnson R. Retrospective assessment of prepubertal major depression with the Kiddie-SADS-E. *Journal of the American Academy of Child Psychiatry* 1982;21:392–397. [PubMed: 7119313]
- Raftery AE. Bayesian model selection in social research. *Sociological Methodology* 1995;25:111–196.
- Rohde P, Lewinsohn PM, Seeley JR, Klein DN, Andrews JA, Small JW. Psychosocial functioning of adults who experienced substance use disorders as adolescents. *Psychology of Addictive Behavior* 2007;21:155–164.
- Sher KJ, Trull TJ. Personality and disinhibitory psychopathology: Alcoholism and antisocial personality disorder. *Journal of Abnormal Psychology* 1994;103:92–102. [PubMed: 8040486]
- Shibata, R. Statistical aspects of model selection. In: Willems, JC., editor. *From data to model*. London: Springer-Verlag; 1989. p. 215–240.
- Slade T, Watson D. The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychological Medicine* 2006;36:1593–1600. [PubMed: 16882356]
- Tackett JL, Krueger RF, Iacono WG, McGue M. Personality in middle childhood: A hierarchical structure and longitudinal connections with personality in late adolescence. *Journal of Research in Personality* 2008;42:1456–1462. [PubMed: 19122851]
- Tsuang MT, Lyons MJ, Meyer JM, Doyle T, Eisen SA, Goldberg J, et al. Co-occurrence of abuse of different drugs in men: The role of drug-specific and shared vulnerabilities. *Archives of General Psychiatry* 1998;55:967–972. [PubMed: 9819064]
- Vollebergh WAM, Iedema J, Bijl RV, de Graaf R, Smit F, Ormel J. The structure and stability of common mental disorders: The NEMESIS study. *Archives of General Psychiatry* 2001;58:597–603. [PubMed: 11386990]
- Waschbusch DA. A meta-analytic examination of comorbid hyperactive–impulsive–attention problems and conduct problems. *Psychological Bulletin* 2002;128:110–150. [PubMed: 11843544]
- Watson D, O'Hara MW, Stuart S. Hierarchical structures of affect and psychopathology and their implications for the classification of emotional disorders. *Depression and Anxiety* 2008;25:282–288. [PubMed: 18415951]
- Wells JE, Horwood LJ. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. *Psychological Medicine* 2004;34:1001–1011. [PubMed: 15554571]
- Wills TA, Dishion TJ. Temperament and adolescent substance use: A transactional analysis of emerging self-control. *Journal of Clinical Child and Adolescent Psychology* 2004;33:69–81. [PubMed: 15028542]
- Windle M, Scheidt DM. Alcoholic subtypes: Are two sufficient? *Addiction* 2004;99:1508–1519. [PubMed: 15585042]

- Wittchen H-U, Höfler M, Merikangas K. Toward the identification of core psychopathological processes? *Archives of General Psychiatry* 1999;56:929–931. [PubMed: 10530635]
- Yang C-C. Evaluating latent class analyses in qualitative phenotype identification. *Computational Statistics & Data Analysis* 2006;50:1090–1104.
- Young SE, Stallings MC, Corley RP, Krauter KS, Hewitt JK. Genetic and environmental influences on behavioral disinhibition. *American Journal of Medical Genetics: Neuropsychological Genetics* 2000;96:684–695.
- Yu, C-Y. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. University of California; Los Angeles: 2002. Unpublished doctoral dissertation

Externalizing Model 1 (EXT1) *Externalizing Model 2a (EXT2a)*

Externalizing Model 2b (EXT2b) *Externalizing Model 3 (EXT3)*

Figure 1.
Pictorial summary of competing theoretical models.

Figure 2.

Path diagram for externalizing model 2b (EXT2b).

Note: Parameter estimates are standardized, and all are significant at $p < .005$.

Table 1

Tetrachoric Correlations (above Diagonal with Standard Errors) and Odds Ratios (below Diagonal with 95% Confidence Intervals) for Externalizing Disorder Diagnostic Categories

	ADHD	ODD	CD	AAB	ALC	CAN	DRG
ADHD	—	0.42 (0.16)	0.34 (0.16)	0.22 (0.17)	0.13 (0.12)	0.16 (0.13)	0.26 (0.13)
ODD	7.6 (2.0–28.9)	—	0.31 (0.16)	0.19 (0.17)	0.21 (0.11)	0.27 (0.12)	0.35 (0.12)
CD	5.0 (1.3–18.6)	4.2 (1.1–15.4)	—	0.67 (0.08)	0.54 (0.08)	0.60 (0.08)	0.62 (0.08)
AAB	2.9 (0.7–12.7)	2.5 (0.6–10.6)	18.5 (8.0–42.5)	—	0.64 (0.07)	0.64 (0.07)	0.47 (0.09)
ALC	1.6 (0.7–4.1)	2.2 (0.9–5.2)	9.0 (3.6–22.6)	15.6 (5.6–43.2)	—	0.64 (0.04)	0.58 (0.06)
CAN	1.9 (0.7–5.2)	2.9 (1.2–7.1)	11.3 (5.1–24.8)	13.0 (6.1–27.5)	7.9 (5.3–11.7)	—	0.77 (0.04)
DRG	3.1 (1.1–8.7)	4.3 (1.7–10.7)	12.2 (5.7–25.7)	6.1 (3.0–12.4)	7.3 (4.4–12.2)	19.4 (11.6–32.4)	—

Note. ADHD = attention-deficit/hyperactivity disorder, ODD = oppositional defiant disorder, CD = conduct disorder, AAB = adult antisocial behavior, ALC = alcohol use disorder, CAN = cannabis use disorder, and DRG = hard drug use disorder.

Table 2

Externalizing Disorders: Model Fit Indices and Selection Criteria

Model	Indicators of Model Fit					Model Selection Criteria				
	χ^2 value	χ^2 df	χ^2 <i>p</i> -value	CFI	TLI	RMSEA	WRMR	Adjusted BIC	AIC	Δ AIC
EXT1	19.35	10	0.036	0.988	0.985	0.034	0.814	−15.94	47.35	2.37
EXT2a	17.22	10	0.070	0.991	0.988	0.030	0.731	−18.07	47.22	2.24
EXT2b	14.98	10	0.133	0.994	0.992	0.025	0.672	−20.31	44.98	0.00
EXT3	13.49	9	0.142	0.994	0.992	0.025	0.586	−18.27	47.49	2.51

Note. CFI = comparative fit index; TLI = Tucker Lewis fit index; RMSEA = root mean square error of approximation; WRMR = root mean square residual; BIC = Bayesian information criterion; AIC = Akaike information criterion. The “best” of the competing models as implied by the AIC statistic is indicated by a value of 0.00 under the Δ AIC column.

Table 3

Psychosocial Outcomes Associated with Model EXT2b

Outcome	Model EXT2b			
	Oppositional Behavior Disorders		Social Norm Violation Disorders	
	Coefficient	<i>p</i>	Coefficient	<i>p</i>
Years of school	–.369	.002	–.115	.249
Unemployment in past year	.021	.849	.120	.102
Income (past year)	–.048	.719	–.325	<.001
Never married	–.327	.086	.347	.011
Currently married	.190	.277	–.405	<.001
Ever divorced/separated	.132	.428	.140	.188
Ever have a biological child	.034	.818	–.020	.834
Relationship quality	–.131	.239	–.143	.069
Poor social adjustment	.059	.631	.190	.012
Poor physical health (past year)	.115	.341	.157	.036
Risky sexual behavior (past year)	–.358	.070	.442	.001
Suicide attempt (since T ₃)	–.020	.909	.434	<.001
Life dissatisfaction	–.023	.837	.214	.004
Coping skills and resources	.147	.205	–.319	<.001
Stressful life events (past year)	.066	.425	.261	<.001
Lifetime depressive disorder	.080	.578	.208	.028
Lifetime anxiety disorder	.191	.237	.075	.497

Note. Statistically significant ($p < .05$) standardized regression coefficients are bolded. *Ns* range from 752 to 816 across analyses.