Criterion Validity and Relationships Between Alternative Hierarchical Dimensional Models of General and Specific Psychopathology

Tyler M. Moore
University of Pennsylvania

Antonia N. Kaczkurkin, E. Leighton Durham,
Hee Jung Jeong, Malerie G. McDowell,
and Randolph M. Dupont
Vanderbilt University

Brooks Applegate
Western Michigan University

Jennifer L. Tackett
Northwestern University

Carlos Cardenas-Iniguez, Omid Kardan, Gaby N. Akcelik, Andrew J. Stier, Monica D. Rosenberg,
Donald Hedeker, Marc G. Berman, and Benjamin B. Lahey
University of Chicago

Psychopathology can be viewed as a hierarchy of correlated dimensions. Many studies have supported this conceptualization, but they have used alternative statistical models with differing interpretations. In bifactor models, every symptom loads on both the general factor and 1 specific factor (e.g., internalizing), which partitions the total explained variance in each symptom between these orthogonal factors. In second-order models, symptoms load on one of several correlated lower-order factors. These lower-order factors load on a second-order general factor, which is defined by the variance shared by the lower-order factors. Thus, the factors in second-order models are not orthogonal. Choosing between these valid statistical models depends on the hypothesis being tested. Because bifactor models define orthogonal phenotypes with distinct sources of variance, they are optimal for studies of shared and unique associations of the dimensions of psychopathology with external variables putatively relevant to etiology and mechanisms. Concerns have been raised, however, about the reliability of the orthogonal specific factors in bifactor models. We evaluated this concern using parent symptom ratings of 9–10 year olds in the ABCD Study. Psychometric indices indicated that all factors in both bifactor and second-order models exhibited at least adequate construct reliability and estimated replicability. The factors defined in bifactor and second-order models were highly to moderately correlated across models, but have different interpretations. All factors in both models demonstrated significant associations with external criterion variables of theoretical and clinical importance, but the interpretation of such associations in second-order models was ambiguous due to shared variance among factors.

General Scientific Summary
Some investigators have proposed that viewing the correlated symptoms of psychopathology as a hierarchy in which all symptoms are related to both a general (*p*) factor of psychopathology and a more specific factor will make it easier to distinguish potential risk factors and mechanisms that are nonspecifically related to all forms of psychopathology versus those that are associated with specific

This article was published Online First July 16, 2020.

Tyler M. Moore, Department of Psychiatry, University of Pennsylvania; Antonia N. Kaczkurkin, E. Leighton Durham, Hee Jung Jeong, Malerie G. McDowell, and Randolph M. Dupont, Department of Psychology, Vanderbilt University; Brooks Applegate, Department of Educational Leadership, Research and Technology, Western Michigan University; Jennifer L. Tackett, Department of Psychology, Northwestern University; Carlos Cardenas-Iniguez, Omid Kardan, and Gaby N. Akcelik, Departments of Psychology and Public Health Sciences, University of Chicago; Andrew J. Stier, Monica D. Rosenberg, Donald Hedeker, and Marc G. Berman, Department of Psychology, University of Chicago; Benjamin B. Lahey, Department of Public Health Sciences (MC 2000), University of Chicago.

dimensions of psychopathology. Parent ratings of child psychopathology items from the Adolescent Brain Cognitive Development (ABCD) Study were analyzed using two alternative statistical models of the proposed hierarchy. All factors of psychopathology defined in both bifactor and second-order models demonstrated adequate psychometric properties and criterion validity, but associations of psychopathology factors with external variables were more easily interpreted in bifactor than in second-order models.

Many scholars argue that psychopathology should not be conceptualized as distinct categorical diagnoses, but as a hierarchy of correlated dimensions of maladaptive behaviors, emotions, and cognitions (Achenbach, Conners, Quay, Verhulst, & Howell, 1989; Helzer, Kraemer, & Krueger, 2006; Krueger et al., 2018; Markon, Chmielewski, & Miller, 2011). Recently, this hierarchy has been posited to consist of a broad general factor of psychopathology—also known as the *p* factor—and two or more specific factors of psychopathology with putatively different causes and mechanisms (Caspi et al., 2014; Caspi & Moffitt, 2018; Kotov et al., 2017; Lahey et al., 2012; Lahey, Krueger, Rathouz, Waldman, & Zald, 2017; Lahey, Van Hulle, Singh, Waldman, & Rathouz, 2011).

One impediment to research on the hierarchical structure of psychopathology is the use of alternative statistical models with different interpretations to define the hierarchies. In the "bassackward" strategy, a series of *separate* correlated-factors analyses are conducted (Goldberg, 2006). One factor is extracted in the first model to define the general factor. Two factors are extracted in a separate second model, followed by three factors in the next model, and so on, until all justifiable factors have been extracted. Because the only connections between levels of such hierarchies are correlations calculated after the fact between factors in the separate successive models, factors are defined in each model without parsing variance attributed to factors at other levels (Michelini et al., 2019).

Some studies use *second-order models* (Carragher et al., 2016; Sunderland et al., 2019), in which every symptom (or dimension of symptoms) loads on one of three or more correlated lower-order factors, which load on the second-order general factor (see Figure 1). Thus, the general factor is defined by the variance shared by the lower-order factors. As a result, the general and lower-order factors in second-order models are not independent, but share variance. Other studies use *bifactor models* (Holzinger & Swineford, 1937; Mansolf & Reise, 2017; Reise, 2012) to define the hierarchy of psychopathology dimensions (Bloemen et al., 2018; Carragher et al., 2016; Caspi et al., 2014; Lahey et al., 2012, 2015, 2018; Sunderland et al., 2019). In bifactor models (see Figure 1), each symptom (or dimension of symptoms) loads both on the general factor and on one of two or more orthogonal specific factors. These loadings completely and optimally partition the total variance in symptoms into separate and orthogonal general and specific factors (Mansolf & Reise, 2017; Reise, 2012).

As detailed in the psychometric literature, bifactor and second-order factor models are similar on the surface, but different in interpretability (Bornovalova, Choate, Fatimah, Petersen, & Wiernik, 2020; Mansolf & Reise, 2017). Although the general factors defined in these two models are highly correlated, even the most straightforward test of criterion validity—the association between *only* the general factor and relevant external variables—has different interpretations in bifactor versus second-order models.

Furthermore, tests of criterion validity of the specific/lower-order factors in these models have different interpretations in two important ways. First, unlike bifactor models, the general factor in a second-order model shares variance with the lower-order factors. This changes the meaning of associations of the general factor with criterion variables. Second, in bifactor models, one can examine the unique associations of external variables with the specific factors while controlling for the general factor (because they are defined independently), but this cannot be done in second-order models because they are not independently defined. This is be-
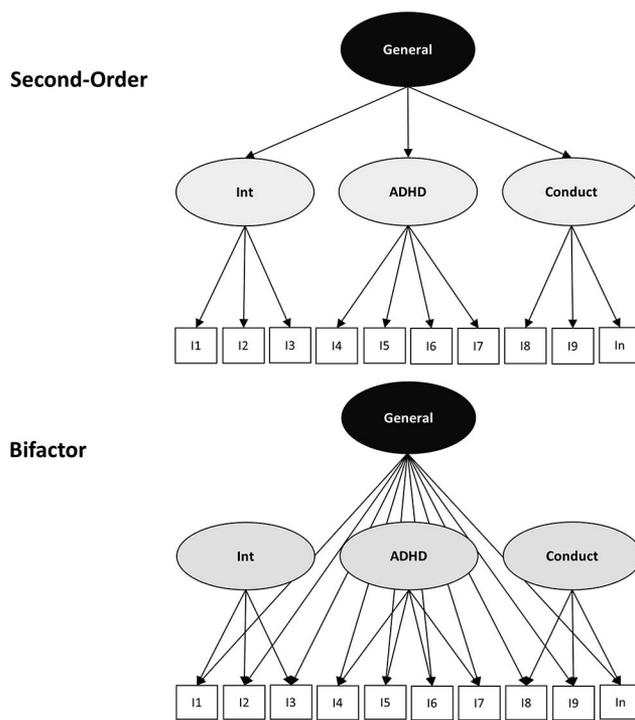
Figure 1. Diagrams of confirmatory factor models compared in the present analyses. ADHD = attention-deficit hyperactivity disorder.

cause the general factor in second-order models is defined solely by the loadings of the lower-order factors on it, making the lower- and second-order general factors (as a whole) *perfectly collinear*. One could identify the external correlates of each lower-level factor one at a time, but the meaning of those associations would be ambiguous as they could partly or wholly reflect variance that is shared with the other factors in the second-order model. Alternatively, associations of each lower-order factor with external variables could be estimated while controlling for the other lower-level factors, but the associations of all three factors with the criteria would likely be overestimated because the lower-order factors would still share variance with general factor in second-order models. If one chose instead to use the residuals ("disturbances") of the lower-order factors (i.e., first regressing out the second-order general factor), 100% of the shared variance between the second-order general factor and each lower-order factor would be attributed to the second-order general factor, underestimating associations of lower-order factors with criteria.

In contrast, because all factors are orthogonal in bifactor models, one can simultaneously regress measured criterion variables on the general and specific factors to determine if each of these factors independently accounts for unique variance in that variable. Such advantages of bifactor models for discovering the correlates of psychopathology at each level of the hierarchy are not in dispute (Bornovalova et al., 2020; Watts, Poore, & Waldman, 2019). What is in dispute, however, is whether the psychometric properties of all factors in bifactor models are sufficient to realize its logical advantages (Conway, Mansolf, & Reise, 2019; Watts et al., 2019). This issue can be addressed in part using psychometric indices recommended for the evaluation bifactor models (Rodriguez, Reise, & Haviland, 2016): (a) H quantifies how well each latent factor is represented by the items loading on it, and is sometimes interpreted as estimating the future replicability of the factor; (b) explained common variance (ECV) is the proportion of the total variance in all items explained by the general factor rather than the specific factors; (c) Omega estimates the proportion of total variance in the symptoms attributable to the general and specific factors together; (d) OmegaH is the proportion of total variance attributable to each general or each specific factor, by itself; (e) factor determinacy estimates the reliability of factor scores from the correlation between a factor and the scores generated from that factor (Grice, 2001); (g) percent uncontaminated correlations (PUC) is the percent of all correlations among symptoms attributable purely to the general factor.

Few published studies of bifactor models have reported these psychometric indices, but two articles (Conway et al., 2019; Watts et al., 2019) have raised concerns about the psychometric adequacy of specific factors in bifactor models based on some H indices below the recommended benchmark of .70 (Hancock & Mueller, 2001). Fortunately, a recent meta-analysis calculated the psychometric indices from factor loading matrices in 49 articles that had not reported them (Constantinou & Fonagy, 2019). This meta-analysis found a mean H for the general factor of 0.91 and 0.69 for specific factors. Mean factor determinacy was adequate to high for all factors, indicating that general and specific factor scores were reliable enough to test their criterion validity. Furthermore, the mean PUC of 0.67 was somewhat below the benchmark of 0.70, thought to indicate that the symptoms are essentially unidimensional, indicating that specific factors could be defined

adequately. Importantly, the psychometric indices were quite heterogeneous in these 49 studies, with these indices varying with the informant on psychopathology and being stronger when more items (i.e., symptoms or dimensions) defined each factor.

In this article, we report the results of new factor analyses of item-level data from parent ratings of child behavior and emotions as well as tests of criterion validity using both models in the Adolescent Brain Cognitive Development (ABCD) Study (Volkow et al., 2018). These head-to-head empirical comparisons are conducted following strict requirements suggested for bifactor modeling (Bonifay, Lane, & Reise, 2017; Bornovalova et al., 2020; Sellbom & Tellegen, 2019):

1. Because bifactor models tend to fit better than correlated-traits and second-order models, even when the data do not justify it (Bonifay et al., 2017), a clear theoretical basis for use of a bifactor model is required. We use bifactor models because they operationalize the causal taxonomy proposed by Lahey, Krueger, Rathouz, Waldman, and Zald (2017), which posits that the hierarchy of correlated phenotypes results from a hierarchy of genetic and environmental causal influences, some of which nonspecifically influence all dimensions of psychopathology to varying extents through the general factor, whereas other causal factors influence dimensions within subdomains of psychopathology (e.g., internalizing and externalizing), and still other causal factors influence only a single dimension of symptoms.

2. Samples must be at least reasonably unbiased and large enough to reliably estimate parameters and psychometric indices.

3. The general and the specific factors defined in bifactor models must be reliably measured and replicable over time in the same persons. The indices just described are useful in this regard, but psychometric indices based on a *single assessment* are not the only, and probably not the best, way to assess the reliability and stability of factors. Notably, several longitudinal studies have found that bifactor structures are replicated and each factor is correlated with itself over time in the same individuals, which could happen only if the factors were reliably measured (Castellanos-Ryan et al., 2016; Gluschkoff, Jokela, & Rosenström, 2019; McElroy, Belsky, Carragher, Fearon, & Patalay, 2018; Neumann et al., 2016; Olino et al., 2018; Snyder, Young, & Hankin, 2017).

4. Each factor must be valid at least in the sense of criterion validity. Notably, the general factor and specific factors of psychopathology defined in bifactor models have each already been found to account for unique variance in adverse functional outcomes, including psychoactive drug prescriptions, incarceration, poor academic progress, suicidal behavior, and self-harm (Haltigan et al., 2018; Lahey et al., 2015; Pettersson, Lahey, Lundström, Larsson, & Lichtenstein, 2018; Sallis et al., 2019), and in theoretically related constructs, including trait negative emotionality (Caspi et al., 2014; Class et al., 2019) and executive functions (Bloemen et al., 2018; Martel et al., 2017; Shields, Reardon,

Brandes, & Tackett, 2019). In the present analyses, we assess the criterion validity of each factor defined in bifactor and second-order models using a diverse set of theoretically and practically relevant external variables that are measured independently (i.e., without shared method variance with psychopathology).

Method

Sample

The present analyses used data from Wave 1 of the ABCD Study. This sample was recruited at 22 sites across the United States at 9–10 years of age as part of a planned longitudinal study. The sites do not represent the population of the United States, but the same unbiased recruitment process was used within every site (Garavan et al., 2018) and poststratification weights can be used to adjust the sample to be more representative (Heeringa & Berglund, 2018). Parent ratings of child psychopathology were collected ($N = 11,866$; 47.9% female). Most (8,142) participants were one child of singleton birth from different families, but 3,724 had a twin or nontwin sibling in the study. Parents classified the children as Non-Hispanic White (52.08%), Black (14.99%), Hispanic (20.29%), and other race-ethnic groups (12.63%). We conducted exploratory and preliminary analyses in one of two stratified random halves of the sample ($N = 5,932$ with nonmissing psychopathology data), selected within data collection sites. The second random half of the sample ($N = 5,934$) was used to conduct planned confirmatory factor analyses that were specified using the results of the exploratory analyses.

Measures

The Child Behavior Checklist (CBCL; Achenbach, 2009) is a parent rating scale of child behavior consisting of 119 items describing behaviors and emotions on a scale of 0 = *not true (as far as you know)*, 1 = *somewhat or sometimes true*, or 2 = *very true or often true*. Missing data on CBCL items was <0.1%. We used other measures administered in the ABCD Study in tests of the criterion validity of psychopathology factors. To avoid confounding by method variance, all criterion measures were based on youth reports or formal testing of the children. Parent reports of functional impairment were the exception, but these represented decisions made by social systems (e.g., special education placement).

Cognitive measures. The ABCD cognitive test battery (Luciana et al., 2018) assessed fluid reasoning, episodic memory, flexible thinking, attention, working memory, learning, mental rotation, processing speed, vocabulary comprehension, and oral reading. It included the Little Man Task (Acker & Acker, 1982), Rey Auditory Verbal Learning Task (mean total correct of Trials 1–5; Taylor, 1959), matrix reasoning subtest from the Wechsler Intelligence Test for Children-V (Wechsler, 2014), Picture Vocabulary Task (PPVT; Gershon et al., 2013), Oral Reading Recognition Task (Gershon et al., 2013), Pattern Comparison Processing Speed Test (Carlozzi, Beaumont, Tulsky, & Gershon, 2015), Dimensional Change Card Sort Task (Zelazo, 2006), List Sorting Working Memory Test (Tulsky et al., 2013), Picture Sequence

Memory Test (Bauer et al., 2013), and Flanker Task (Fan, McCandliss, Sommer, Raz, & Posner, 2002).

Dispositional measures. The ABCD battery (Barch et al., 2018) included two measures of dispositional traits completed by child self-report: an ABCD short form of the UPPS impulsivity measure (Zapolski, Stairs, Settles, Combs, & Smith, 2010) and the prosocial subscale from the Strengths and Difficulties Questionnaire (Goodman, 1997; Hawes et al., 2019).

Suicide and self-harm behavior. Children were administered a version of the Kiddie Schedule for Affective Disorders and Schizophrenia in which they reported if they had ever engaged in self-injurious behavior without suicidal intent, concrete suicidal ideation or planning, an interrupted, aborted, or completed suicide attempt, or planned a suicide attempt (Kaufman et al., 2016; Lisdahl et al., 2018).

Functional impairment. Three measures of functional impairment included ever receiving mental health services, school detentions or suspensions, and enrollment in any form of special education, except gifted programs.

Statistical Analyses

Missing data were minimal; only two of 5,936 selected for the second split sample had missing CBCL data. Among the 5,934 with nonmissing CBCL data, 93.61% had nonmissing data on all cognitive variables. Similarly, 99.66% had nonmissing data on all dispositional scales and 96.66% had nonmissing values on all impairment measures. All factor analyses and structural equation models were conducted in Mplus 8.3 using the mean- and variance-adjusted weighted least squares (WSLMV) estimator, which uses pairwise deletion for missing data (Muthén & Muthén, 1998–2017). All analyses accounted for the stratification of the sample in data collection sites, used poststratification weights (Heeringa & Berglund, 2018), and accounted for clustering within families.

Exploratory Analyses in the First Half of the Sample

Exploratory analyses reduced the number of CBCL items to those most strongly associated with psychopathology at this age. The CBCL casts a broad net, including behaviors of concern to parents not typically viewed as symptoms of psychopathology, such as biting fingernails and constipation. Additionally, the CBCL includes items that are more appropriate for older ages, such as alcohol consumption and smoking. The item "wishes to be of the opposite sex," was eliminated because it does not reflect psychopathology. The steps in the exploratory analyses were: Eight items referring to behaviors typical of adolescents were eliminated because ratings above 0 were <0.5%, or were <1.0% *and* it was not possible to estimate polychoric correlations with other items (Table S1). Three pairs of items that referenced similar behaviors correlated >.85 were combined in composites by taking the mean rating of the items and rounding to achieve 0, 1, 2 scoring (Table S1). Exploratory structural equation (ESEM) models (Asparouhov & Muthén, 2009) were conducted with OBLIMIN rotation. Parallel analysis (Horn, 1965) with Glorfeld correction (Glorfeld, 1995) indicated that up to nine factors could be extracted (Figure S1). The minimum average partial criterion (Velicer, 1976) indicated that up to six factors could be extracted (Figure

S1). We extracted four interpretable factors and all CBCL items with a loading ≥0.40 on at least one factor were retained. Retained items were subjected to ESEM specifying two, three, or four correlated factors in the first half of the sample (Tables S2–S4).

Confirmatory Analyses in the Second Half of the Sample

Bifactor and second-order confirmatory models were specified based on the results of the ESEMs. Items loading >0.40 on two factors were assigned to the factor with the higher loading. As required for bifactor models, each item loaded on the general factor and only one specific factor. All other loadings were fixed to zero, and all factors were specified to be orthogonal (Reise, 2012).

Comparisons of the Models

We directly compared the bifactor and second-order models specifying a general factor and three specific or three lower-order factors of psychopathology in two ways. First, estimated factor scores from Mplus were plotted. Second, tests of criterion validity were conducted using structural equation modeling in Mplus. For the bifactor model, each measured criterion variable was regressed simultaneously in these SEMs on the general and specific psychopathology factors. Probit regressions were estimated in SEM for binary criterion variables. Based on previous studies that identified a general factor of cognitive ability test scores using similar test batteries in children (Friedman & Miyake, 2017; Martel et al., 2017), the substantially correlated cognitive measures defined a single latent factor in a confirmatory measurement model.

The strategy for testing the criterion validity of the second-order model was necessarily different. Because the second-order general factor is defined by the lower-order factors, it was necessary to conduct two separate regression analyses (within the SEMs), one for only the general factor and one in which criteria were simultaneously regressed on lower-order internalizing, conduct problems, and attention-deficit/hyperactivity disorder (ADHD) factors. For both bifactor and second-order models, covariates of no interest (child's age in months, sex, and race-ethnicity) were included in the models.

Sensitivity Analysis

To evaluate the impact of missing data on cognitive test measures, we first used listwise deletion to drop all participants with missing data on any measure. Second, we imputed missing data once for cognitive measures and using SAS PROC MI.

Results

ESEMs Based in the First Half of the Sample

Factor loadings for two-, three-, and four-factor ESEMs of the retained CBCL items are in Tables S2–S4. When three factors were extracted, the items defined factors of internalizing (i.e., fears, depression, insecurity, and somatic complaints), conduct problems (i.e., oppositional and conduct disorder), and ADHD and neurodevelopmental problems (e.g., immaturity, poor coordina-

tion, and sluggish cognitive tempo), all of which have long been known to be correlated (Achenbach et al., 1989; Quay, 1986). Items referring to strange ideas and behaviors loaded on the ADHD factor. When a fourth factor was extracted, somatic complaint items split from the internalizing factor.

Confirmatory Factor Analyses in the Second Half of the Sample

Results for a bifactor model with two specific factors are in Table S5. We focused on bifactor and second-order models with three or four specific/lower-order factors, however, which allow second-order models to be identified. Factor loadings from models with three specific/lower-order factors are in Tables S6 and S7 (Tables S8 and S9 for models specifying four specific/lower-order factors). Each model fit the data well (Table S10). Because of difficulties in choosing between substantively different models that fit the data well using fit statistics, we evaluated the two structural models in terms of their criterion validity (Bonifay et al., 2017; Greene et al., 2019).

As shown in Table 1, all factors in both models had acceptable H indices >0.70 (Hancock & Mueller, 2001) and each specific factor in the bifactor models was reliable according to omega statistics. ECV and OmegaH indicated that the general factor is robust and explains the majority of the estimated reliability of each specific factor. Nonetheless, the factor determinacy indices show that each of the specific factors is sufficiently determined to calculate specific factor scores that are orthogonal to each other and to the general factor score. This means that any unique associations of a specific factor score with external variables can be interpreted as unique to that factor and can be used to evaluate their criterion validity and examine associations with variables relevant to causes and mechanisms.

Correlations Between Factors Across Models

As shown in Figure 2, estimated factor scores for the general factors defined in the two models were highly correlated. In contrast, the specific internalizing, conduct, and ADHD factor scores were moderately correlated across models. The relationships in Figure 2 are more complex than they first appear, however. Because general and specific/lower-order factor scores are independent in bifactor models, as shown in Figures S3 and S4, individuals high on ADHD, for example, are often low on the general factor and vice versa. In contrast, that is not the case in second-order models because these factors are highly correlated. As a result, Figure 2 reveals an apparent Simpson's Paradox (Simpson, 1951) in which the overall *positive linear* correlations between specific and lower-order scores across models obscure the fact that, at low values, ADHD scores, for example, from the two models are *negatively correlated*. This occurs because at higher general factor scores, ADHD scores in both models tend to be high, but when general factor scores are low, ADHD scores are lower in second-order than bifactor models. This negative correlation in the lower range is one of the ways the bifactor model achieves specific factor scores orthogonal to the general factor when considered across the full range of scores.

Table 1

Psychometric Indices for Each Latent Factor Defined in Bifactor and Second-Order Models Specifying Three or Four Specific/Lower-Level in the Second Split Half of the ABCD Study Sample (N = 5,926) to Which the Statistic Is Applicable

Factors	General	Externalizing	ADHD	Internalizing	Somatic
Index	Bifactor model with three specific factors				
H	0.979	0.860	0.730	0.847	
ECV	0.703	0.127	0.059	0.111	
Omega	0.983	0.982	0.981	0.982	
OmegaH	0.872	0.060	0.017	0.042	
Factor determinacy	0.978	0.930	0.897	0.925	
PUC	0.640				
	Bifactor model with four specific factors				
H	0.980	0.867	0.725	0.787	0.808
ECV	0.693	0.122	0.053	0.072	0.060
Omega	0.984	0.984	0.983	0.983	0.983
OmegaH	0.889	0.058	0.016	0.021	0.008
Factor determinacy	0.977	0.930	0.891	0.898	0.927
PUC	0.694				
	Second-order model based on three lower-level factors				
H	—	0.976	0.953	0.949	
Factor determinacy	0.942	0.988	0.978	0.975	
	Second-order model based on four lower-level factors				
H	—	0.976	0.955	0.944	0.907
Factor determinacy	0.947	0.988	0.979	0.974	0.955

Note. ADHD = attention-deficit hyperactivity disorder/sluggish cognitive tempo; H = index of construct replicability; ECV = explained common variance; PUC = percent correlations that are uncontaminated; OmegaH = ω-hierarchical. Not all indices are applicable to the second-order factor model.

Criterion Validity

Simultaneous regressions within each structural equation model (SEM) revealed that each of the latent general and specific factors defined by the bifactor model explained significant unique variance after correction for multiple testing in the criterion variables (see Table 2). The latent second-order general factor was significantly associated with the same criterion variables as in the bifactor models (see Table 3) and each of the three lower-level factors explained significant unique variance in multiple criterion variables, controlling for the variance explained by the other two lower-order factors. The results for the bifactor model with four specific factors and the second-order model based on four lower-order factors presented in Tables S11 and S12 were virtually identical, except for the addition of the somatic complaints factor. Sensitivity analyses based on listwise deletion and imputation of missing cognitive test scores yielded similar results (Table S13).

Discussion

This study provides new information on two alternative statistical models of the hierarchical dimensional structure of psychopathology using data from the largest study of child psychopathology to date. Both bifactor and second-order factor models of CBCL symptoms fit the data well. The psychometric indices reported in Table 1 provide strong support for the general factor in both models; indeed, they raise the radical possibility that psychopathology measured by the CBCL at this age is unidimensional (i.e., adequately captured by the general factor alone). The OmegaH (~0.87) and ECV (~0.70) indicated that the general factor is strong, particularly when considered with the PUC, which indicates that 64% of item correlations are explained by the general

factor. Furthermore, when omega and OmegaH are compared, >90% of the reliability of the specific factor scores (e.g., specific conduct problems) is due to the general factor. It should be noted, however, that a study of psychopathology in older youth (Sunderland et al., 2019) found far less evidence of unidimensionality.

Moreover, other results of the present analyses argue against unidimensionality and support the *multidimensionality* of parent-rated psychopathology. Like Sunderland et al. (2019), the present bifactor and second-order models both defined all factors with strong determinacy and acceptable construct replicability (see Table 1). Therefore, it was possible to conduct critical tests of the criterion validity to determine if the lower-level/specific factors are valid.

Criterion Validity of General and Specific/Lower-Order Factors

Like previous studies (Lahey et al., 2017; Pettersson et al., 2018), the general factor defined in the bifactor model accounted for substantial unique variance in external criteria. For example, as presented in Table 2, each 1 *SD* increase in general factor scores was associated with a .38 *SD* (± 0.03) greater likelihood of being detained or suspended from school for misbehavior, over and above demographic covariates and the robust association with specific conduct problems and the inverse association with internalizing psychopathology. Furthermore, *only* the general factor was significantly correlated with youth-reported suicidal planning and attempts, and each 1 *SD* increase in general factor scores was associated with a .27 *SD* (± 0.04) greater likelihood of engaging in nonsuicidal self-injury. The general factor in the bifactor model
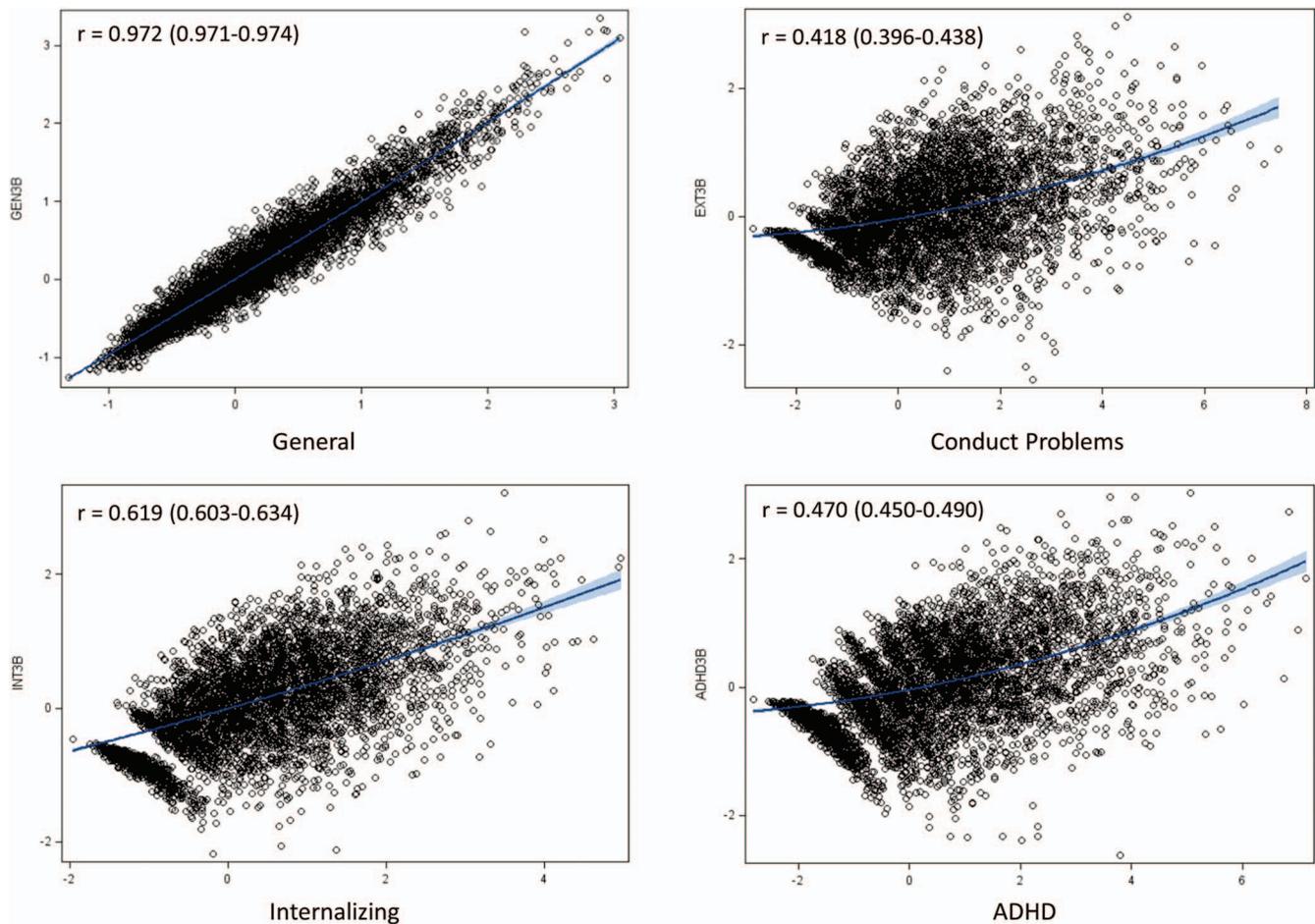
Figure 2. Scatterplots and correlations between estimated general and specific or lower-order factor scores defined in bifactor (vertical axes) and second-order confirmatory factor models (horizontal). Blue shading defines the 95% confidence interval of the best-fit regression line. ADHD = attention-deficit hyperactivity disorder. See the online article for the color version of this figure.

also was significantly associated with obtaining mental health and special education services. These associations suggest the potential clinical utility of the general factor in evaluating prognostic risk in children.

At a theoretical level, it is important that the results in Table 2 replicate previous associations of the bifactor general factor with independent measures of cognitive functioning (Martel et al., 2017; Shields et al., 2019) and negative emotionality (Caspi & Moffitt, 2018; Lahey et al., 2017). Although negative emotionality was not measured in this study, UPPS negative urgency was measured, which is correlated with negative emotionality (Settles et al., 2012). Thus, the present findings that the general factor was significantly associated with both cognitive functioning and UPPS negative urgency are consistent with the hypothesis that lower executive functioning and higher negative emotionality are two processes that nonspecifically contribute to risk for all forms of psychopathology, which is captured by the general factor (Lahey et al., 2017). Additionally, the finding that both positive and negative urgency are associated with the general factor in the same direction is

consistent with the hypothesis that the general factor reflects impulsive (i.e., unregulated) responsivity to both positive and negative emotions (Carver, Johnson, & Timpano, 2017).

Associations of the specific factors defined in the bifactor model with criterion variables reported in Table 2 replicate previous studies (reviewed by Lahey et al., 2017). Consistent with the meanings of these factors, the specific internalizing and conduct problems factors were each positively associated with mental health service, but associated with school detentions and expulsions, and with premeditation, negative and positive urgency, and sensation seeking in opposite directions as would be expected. The specific ADHD factor exhibited criterion validity in its associations with risk for special education placement for learning problems and nonsuicidal self-harm, with three dispositional constructs theoretically associated with ADHD (low premeditation, low perseverance, and high positive urgency), and in its robust inverse association with cognitive functioning. Considered together, these robust associations with independently measured criterion variables support the validity of both the general and the three specific factors defined in the bifactor model.

Table 2

Results of Simultaneous Regressions (Within an SEM) of Each Independently Measured Criterion Variable on the Latent General Factor and Three Specific Factors Defined in a Bifactor Model and on Demographic Covariates of No Interest (Age, Sex, and Race-Ethnicity) in the Second Split Half of the ABCD Study Sample (N = 5,926)

	Bifactor model with three specific factors							
	Factors							
	General		Specific internalizing		Specific conduct		Specific ADHD	
Criterion variable	β	$p <$	β	$p <$	β	$p <$	β	$p <$
Functional impairment (binary)								
Detention/suspension	**0.380**	**0.001**	**−0.084**	**0.011**	**0.409**	**0.001**	0.040	0.307
Mental health service	**0.577**	**0.001**	**0.131**	**0.001**	**0.106**	**0.001**	0.064	0.040
Special education (behavior)	**0.295**	**0.001**	0.024	0.704	**0.253**	**0.001**	0.093	0.112
Special education (learning)	**0.288**	**0.001**	0.003	0.956	0.059	0.212	**0.187**	**0.004**
Youth-reported harmful behaviors (binary)								
Suicidal behavior	**0.270**	**0.001**	0.032	0.456	0.098	0.072	−0.002	0.979
Nonsuicidal self-harm	**0.220**	**0.001**	0.057	0.098	0.081	0.076	**0.093**	**0.025**
Youth-rated dispositional constructs (standardized continuous)								
UPPS low premeditation	**0.132**	**0.001**	**−0.079**	**0.001**	**0.099**	**0.001**	**0.109**	**0.001**
UPPS low perseverance	**0.128**	**0.001**	0.018	0.310	**0.075**	**0.001**	**0.248**	**0.001**
UPPS negative urgency	**0.172**	**0.001**	**−0.059**	**0.002**	**0.087**	**0.001**	−0.001	0.966
UPPs positive urgency	**0.113**	**0.001**	**−0.089**	**0.001**	**0.097**	**0.001**	**0.085**	**0.001**
UPPS sensation seeking	−0.005	0.796	**−0.051**	**0.007**	**0.104**	**0.001**	0.034	0.169
Prosociality	**−0.073**	**0.001**	0.033	0.088	**−0.065**	**0.008**	−0.030	0.211
Latent cognitive test scores								
Cognitive functioning	**−0.091**	**0.001**	0.021	0.293	**−0.241**	**0.001**	**−0.328**	**0.001**

Note. Coefficients in bold are significant after FDR correction (adopting a 5% false discovery rate) for 52 tests. FDR = false discovery rate; ADHD = attention-deficit hyperactivity disorder/sluggish cognitive tempo. Continuous dispositional and cognitive functioning measures were normalized to a mean of 0 and standard deviation of 1.

The results in Table 3 similarly support the criterion validity of psychopathology factors defined in the second-order model. The *pattern* of significant and nonsignificant associations is similar, but the *interpretations* of these associations are different. Because the second-order general factor is defined by the loadings of lower-order factors on it, it was not possible to simultaneously regress criterion variables on both the general and the lower-order factors to test each factor's unique criterion validity. Therefore, tests of criterion validity of lower-order factors and the second-order general factor had to be conducted separately—identifying the unique correlates of the lower-level factors while controlling for the second-order general factor is impossible. It would be analogous to including three items in a predictive model along with the sum of those same three items.

Notably, the latent general factor defined in the second-order model explained an average of 4.9% more variance in these criterion variables (range 1.2–11.5%; median 4.3%) than the general factor in the bifactor model. This does not mean that the second-order model is superior to the bifactor model, however. Indeed, the apparent superiority of the second-order model *reflects its major limitation* for research on variables related to the causes and mechanisms of each factor of psychopathology. The stronger associations of the second-order general factor with criterion variables are similarly due to *contamination* of the general factor and lower-order factors with each other in the second-order model. For example, the association of the second-order general factor with school detentions/suspensions is inflated because the general factor contains variance associated with the lower-order conduct problems factor, which is strongly associated with detentions/suspensions. We can see this at the top of Table 2 where the

variance explained in detentions/suspensions by the general and specific factors in the bifactor model is $0.380^2 + (−0.084)^2 + 0.409^2 + 0.040^2 = 0.32$, with half of that variance coming from conduct problems (0.409). The variance associated with conduct problems is not included in the estimated association of the general factor with detentions/suspensions in the bifactor model, but it spuriously boosts the variance explained by the general factor in the second-order model. This difference in interpretation of such associations is a key difference between the models.

Choosing Between Bifactor and Higher-Order Models of Psychopathology

Given that all latent factors defined in both bifactor and second-order models were well-defined and valid, the choice between these two models depends on its intended use. The bifactor model is useful for testing models that posit that some risk factors and psychobiological mechanisms are common to all forms of psychopathology, whereas other risk factors are only related to specific factors (Caspi et al., 2014; Caspi & Moffitt, 2018; Lahey et al., 2011; Lahey et al., 2017). This is because, in a bifactor model, the variance in symptoms is empirically partitioned among the general and specific factors (Mansolf & Reise, 2017; Reise, 2012). Thus, the bifactor model is useful when the goal is to identify both common and *unique* correlates of the dimensions of psychopathology.

Limitations and Implications for Future Research

It should be emphasized that the results of any factor analysis reflect the items that are analyzed and the sample in which they are

Table 3

Results of Regressions of Each Independently Measured Criterion Variable on Only the Latent General Factor and Demographic Covariates of No Interest (Age, Sex, and Race–Ethnicity) and in Separate Models Simultaneously on Only the Three Lower-Order Factors Defined in Second-Order Factor Models and on Demographic Covariates of No Interest (Age, Sex, and Race–Ethnicity) in the Second Split Half of the ABCD Study Sample (N = 5,926)

			Factors					
	General		Lower-order internalizing		Lower-order conduct		Lower-order ADHD	
Criterion variable	β	p <	β	p <	β	p <	β	p <
Second-order model based on three lower-order factors								
Regressions of criterion variables on only the second-order general factor and covariates								
Functional impairment (binary)								
Detention/suspension	**0.495**	**0.001**						
Mental health service	**0.634**	**0.001**						
Special educ (behavior)	**0.381**	**0.001**						
Special educ (learning)	**0.332**	**0.001**						
Youth-reported harmful behaviors (binary)								
Suicidal behavior	**0.301**	**0.001**						
Nonsuicidal self-harm	**0.263**	**0.001**						
Youth-rated dispositional constructs (standardized continuous)								
UPPS low premeditation	**0.161**	**0.001**						
UPPS low perseverance	**0.188**	**0.001**						
UPPS negative urgency	**0.184**	**0.001**						
UPPS positive urgency	**0.135**	**0.001**						
UPPS sensation seeking	0.017	0.344						
Prosociality	**−0.088**	**0.001**						
Latent general cognitive functioning test scores								
Cognitive functioning	**−0.197**	**0.001**						
Simultaneous regressions of criterion variables on only the three lower-order factors and covariates								
Functional impairment (binary)								
Detention/suspension			**−0.218**	**0.001**	**0.683**	**0.001**	−0.020	0.705
Mental health service			**0.247**	**0.001**	**0.251**	**0.001**	**0.170**	**0.001**
Special Educ (behavior)			−0.027	0.762	**0.372**	**0.001**	0.027	0.748
Special Educ (learning)			−0.017	0.843	0.028	0.758	**0.332**	**0.001**
Youth-reported harmful behaviors (binary)								
Suicidal behavior			0.083	0.245	**0.174**	**0.032**	0.053	0.498
Nonsuicidal self-harm			0.089	0.101	0.063	0.351	0.125	0.044
Youth-rated dispositional constructs (standardized continuous)								
UPPS low premeditation			**−0.143**	**0.001**	**0.091**	**0.001**	**0.203**	**0.001**
UPPS low perseverance			−0.052	0.059	**−0.098**	**0.002**	**0.346**	**0.001**
UPPS negative urgency			**−0.070**	**0.011**	**0.211**	**0.001**	0.034	0.312
UPPs positive urgency			**−0.168**	**0.001**	**0.137**	**0.001**	**0.151**	**0.001**
UPPS sensation seeking			**−0.125**	**0.001**	**0.109**	**0.002**	0.018	0.614
Positive emotionality			**−0.093**	**0.007**	0.048	0.092	−0.038	0.269
Latent general cognitive functioning test scores								
Cognitive functioning			**0.209**	**0.001**	0.065	0.052	**−0.327**	**0.001**

Note. Coefficients in bold are significant after FDR correction (adopting a 5% false discovery rate) for 52 tests. FDR = false discovery rate; ADHD = attention-deficit hyperactivity disorder/sluggish cognitive tempo. Continuous dispositional and executive functioning measures were normalized to a mean of 0 and standard deviation of 1.

analyzed. We restate this obvious fact to guard against the reification of constructs such as the general factor of psychopathology. This admonition is not unique to bifactor models, but applies to any factor analysis. There are several important implications of the present findings for future studies, especially in the context of previous research. These findings indicate that the factors defined in the bifactor model based on CBCL items have sufficient psychometric properties (e.g., determinacy) to allow the advantages of the orthogonal factors in bifactor models to be realized in discovery studies, including the important longitudinal ABCD Study. This means, for example, that the bifactor model can be used to test associations of orthogonal factors of psychopathology to polygenic risk scores in this sample. The general factor has already been shown to be moderately heritable and significantly related to such independently measured genetic risk scores (Allegrini et al., 2020; Neumann et al., 2016), but we need to learn much more. Similarly, well-powered analyses of ABCD structural and functional neuroimaging data will not only help reveal important psychobiological mechanisms related to psychopathology risk in specific and transdiagnostic ways, but could further support the validity of factor scores defined in bifactor models (Bornovalova et al., 2020).

Note that the present findings support such research in the ABCD Study, but do not mean that bifactor models of psychopathology will always have adequate psychometric properties; as with any factor model, that needs to be demonstrated on a study by study basis. Adequate psychometric properties flow from adequate methodology. In a broader sense, the most important obstacle to the field is that no study of the hierarchy of psychopathology has used an instrument that comprehensively measures all symptoms of all forms of psychopathology, including what were previously referred to as clinical and personality disorders (Forbes et al., 2017). If such a comprehensive measure is successfully developed and validated, we will finally be in a position to define the hierarchical structure of psychopathology and to discover its correlates and causes in future studies. One very helpful but inexpensive step toward this goal would be to supplement the CBCL with additional items in future waves of the ABCD Study.

It is important to note the implications of one current finding for the measurement of psychopathology in future waves of the ABCD Study using the CBCL. It was not possible to include a number of CBCL items in the present analyses due to their low endorsements at ages 9–10 years. These items refer to symptoms that are centrally important to the ABCD Study such as psychotic symptoms, drinking, and substance use, but are uncommon at this age. As the ABCD sample ages, however, they will become more prevalent. This means that it likely will not be possible to base factor analyses on the same subset of items at different ages. A valid way to characterize both stability and developmental change in psychopathology in the ABCD sample must be developed.

References

Achenbach, T. M. (2009). *The Achenbach System of Empirically Based Assessment (ASEBA): Development, findings, theory, and applications.* Burlington, VT: University of Vermont Research Center for Children, Youth and Families.

Achenbach, T. M., Conners, C. K., Quay, H. C., Verhulst, F. C., & Howell, C. T. (1989). Replication of empirically derived syndromes as a basis for taxonomy of child/adolescent psychopathology. *Journal of Abnormal Child Psychology, 17,* 299–323. http://dx.doi.org/10.1007/BF00917401

Acker, W., & Acker, W. (1982). *Bexley Maudsley automated processing screening and Bexley Maudsley category sorting test manual.* London, UK: NFER-Nelson Publishing.

Allegrini, A. G., Cheesman, R., Rimfeld, K., Selzam, S., Pingault, J. B., Eley, T. C., & Plomin, R. (2020). The p factor: Genetic analyses support a general dimension of psychopathology in childhood and adolescence. *Journal of Child Psychology and Psychiatry, 61,* 30–39. http://dx.doi.org/10.1111/jcpp.13113

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 26,* 397–438. http://dx.doi.org/10.1080/10705510903008204

Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., . . . Sher, K. J. (2018). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Developmental Cognitive Neuroscience, 32,* 55–66. http://dx.doi.org/10.1016/j.dcn.2017.10.010

Bauer, P. J., Dikmen, S. S., Heaton, R. K., Mungas, D., Slotkin, J., & Beaumont, J. L. (2013). III. NIH Toolbox Cognition Battery (CB): Measuring episodic memory. *Monographs of the Society for Research in Child Development, 78,* 34–48. http://dx.doi.org/10.1111/mono.12033

Bloemen, A. J. P., Oldehinkel, A. J., Laceulle, O. M., Ormel, J., Rommelse, N. N. J., & Hartman, C. A. (2018). The association between executive functioning and psychopathology: General or specific? *Psychological Medicine, 48,* 1787–1794. http://dx.doi.org/10.1017/S0033291717003269

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science, 5,* 184–186. http://dx.doi.org/10.1177/2167702616657069

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biological Psychiatry.* Advance online publication. http://dx.doi.org/10.1016/j.biopsych.2020.01.013

Carlozzi, N. E., Beaumont, J. L., Tulsky, D. S., & Gershon, R. C. (2015). The NIH Toolbox pattern comparison processing speed test: Normative data. *Archives of Clinical Neuropsychology, 30,* 359–368. http://dx.doi.org/10.1093/arclin/acv031

Carragher, N., Teesson, M., Sunderland, M., Newton, N. C., Krueger, R. F., Conrod, P. J., . . . Slade, T. (2016). The structure of adolescent psychopathology: A symptom-level analysis. *Psychological Medicine, 46,* 981–994. http://dx.doi.org/10.1017/S0033291715002470

Carver, C. S., Johnson, S. L., & Timpano, K. R. (2017). Toward a functional view of the p factor in psychopathology. *Psychological Science, 5,* 880–889. http://dx.doi.org/10.1177/2167702617710037

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., . . . Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2,* 119–137. http://dx.doi.org/10.1177/2167702613497473

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *The American Journal of Psychiatry, 175,* 831–844. http://dx.doi.org/10.1176/appi.ajp.2018.17121383

Castellanos-Ryan, N., Brière, F. N., O'Leary-Barrett, M., Banaschewski, T., Bokde, A., Bromberg, U., . . . Conrod, P. (2016). The structure of psychopathology in adolescence and its common personality and cognitive correlates. *Journal of Abnormal Psychology, 125,* 1039–1052. http://dx.doi.org/10.1037/abn0000193

Class, Q. A., Rathouz, P. J., Van Hulle, C. A., Applegate, B., Waldman, I. D., Zald, D. H., & Lahey, B. B. (2019). Socioemotional dispositions of children and adolescents predict general and specific second-order factors of psychopathology in early adulthood across informants: A

12-year prospective study. *Journal of Abnormal Psychology, 128,* 574–584. http://dx.doi.org/10.1037/abn0000433

Constantinou, M., & Fonagy, P. (2019). Evaluating bifactor models of psychopathology using model-based reliability indices. *PsyArXiv.* http://dx.doi.org/10.31234/osf.io/6tf7j

Conway, C. C., Mansolf, M., & Reise, S. P. (2019). Ecological validity of a quantitative classification system for mental illness in treatment-seeking adults. *Psychological Assessment, 31,* 730–740. http://dx.doi.org/10.1037/pas0000695

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14,* 340–347. http://dx.doi.org/10.1162/089892902317361886

Forbes, M. K., Kotov, R., Ruggero, C. J., Watson, D., Zimmerman, M., & Krueger, R. F. (2017). Delineating the joint hierarchical structure of clinical and personality disorders in an outpatient psychiatric sample. *Comprehensive Psychiatry, 79,* 19–30. http://dx.doi.org/10.1016/j.comppsych.2017.04.006

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86,* 186–204. http://dx.doi.org/10.1016/j.cortex.2016.04.023

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., . . . Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience, 32,* 16–22. http://dx.doi.org/10.1016/j.dcn.2018.04.004

Gershon, R. C., Slotkin, J., Manly, J. J., Blitz, D. L., Beaumont, J. L., Schnipke, D., . . . Weintraub, S. (2013). IV. NIH Toolbox Cognition Battery (CB): Measuring language (vocabulary comprehension and reading decoding). *Monographs of the Society for Research in Child Development, 78,* 49–69. http://dx.doi.org/10.1111/mono.12034

Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55,* 377–393. http://dx.doi.org/10.1177/0013164495055003002

Gluschkoff, K., Jokela, M., & Rosenström, T. (2019). The general psychopathology factor: Structural stability and generalizability to within-individual changes. *Frontiers in Psychiatry, 10,* 594. http://dx.doi.org/10.3389/fpsyt.2019.00594

Goldberg, L. R. (2006). Doing it all bass-ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality, 40,* 347–358. http://dx.doi.org/10.1016/j.jrp.2006.01.001

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38,* 581–586. http://dx.doi.org/10.1111/j.1469-7610.1997.tb01545.x

Greene, A. L., Eaton, N. R., Li, K., Forbes, M. K., Krueger, R. F., Markon, K. E., . . . Kotov, R. (2019). Are fit indices used to test psychopathology structure biased? A simulation study. *Journal of Abnormal Psychology, 128,* 740–764. http://dx.doi.org/10.1037/abn0000434

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6,* 430–450. http://dx.doi.org/10.1037/1082-989X.6.4.430

Haltigan, J. D., Aitken, M., Skilling, T., Henderson, J., Hawke, L., Battaglia, M., . . . Andrade, B. F. (2018). "P" and "DP:" Examining symptom-level bifactor models of psychopathology and dysregulation in clinically referred children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 57,* 384–396. http://dx.doi.org/10.1016/j.jaac.2018.03.010

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In S. T. R. Cudeck & D. Sörbom (Eds.), *Structural equation modeling: Present and future—A Festschrift in honor of Karl Jöreskog* (pp. 195–216). Lincolnwood, IL: Scientific Software International.

Hawes, S. W., Waller, R., Thompson, W. K., Hyde, L. W., Byrd, A. L., Burt, S. A., . . . Gonzalez, R. (2019). Assessing callous-unemotional traits: Development of a brief, reliable measure in a large and diverse sample of preadolescent youth. *Psychological Medicine, 50,* 1–9. http://dx.doi.org/10.1017/S0033291719002903

Heeringa, S. G., & Berglund, P. A. (2018). *A guide for population-based analysis of the Adolescent Brain Cognitive Development (ABCD) Study baseline data.* Ann Arbor, MI: Institute for Social Research.

Helzer, J. E., Kraemer, H. C., & Krueger, R. F. (2006). The feasibility and need for dimensional psychiatric diagnoses. *Psychological Medicine, 36,* 1671–1680. http://dx.doi.org/10.1017/S003329170600821X

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2,* 41–54. http://dx.doi.org/10.1007/BF02287965

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185. http://dx.doi.org/10.1007/BF02289447

Kaufman, J., Birmaher, B., Axelson, D., Perepletchokova, F., Brent, D., & Ryan, N. (2016). *Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL-5).* Baltimore, MD: Johns Hopkins University.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126,* 454–477. http://dx.doi.org/10.1037/abn0000258

Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., . . . Zimmermann, J. (2018). Progress in achieving quantitative classification of psychopathology. *World Psychiatry, 17,* 282–293. http://dx.doi.org/10.1002/wps.20566

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology, 121,* 971–977. http://dx.doi.org/10.1037/a0028355

Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin, 143,* 142–186. http://dx.doi.org/10.1037/bul0000069

Lahey, B. B., Rathouz, P. J., Keenan, K., Stepp, S. D., Loeber, R., & Hipwell, A. E. (2015). Criterion validity of the general factor of psychopathology in a prospective study of girls. *Journal of Child Psychology and Psychiatry, 56,* 415–422. http://dx.doi.org/10.1111/jcpp.12300

Lahey, B. B., Van Hulle, C. A., Singh, A. L., Waldman, I. D., & Rathouz, P. J. (2011). Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Archives of General Psychiatry, 68,* 181–189. http://dx.doi.org/10.1001/archgenpsychiatry.2010.192

Lahey, B. B., Zald, D. H., Perkins, S. F., Villalta-Gil, V., Werts, K. B., Van Hulle, C. A., . . . Waldman, I. D. (2018). Measuring the hierarchical general factor model of psychopathology in young adults. *International Journal of Methods in Psychiatric Research, 27,* e1593. http://dx.doi.org/10.1002/mpr.1593

Lisdahl, K. M., Sher, K. J., Conway, K. P., Gonzalez, R., Feldstein Ewing, S. W., Nixon, S. J., . . . Heitzeg, M. (2018). Adolescent brain cognitive development (ABCD) study: Overview of substance use assessment methods. *Developmental Cognitive Neuroscience, 32,* 80–96. http://dx.doi.org/10.1016/j.dcn.2018.02.007

Luciana, M., Bjork, J. M., Nagel, B. J., Barch, D. M., Gonzalez, R., Nixon, S. J., & Banich, M. T. (2018). Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Developmental Cognitive Neuroscience, 32,* 67–79. http://dx.doi.org/10.1016/j.dcn.2018.02.006

Mansolf, M., & Reise, S. P. (2017). When and why the second-order and bifactor models are distinguishable. *Intelligence, 61,* 120–129. http://dx.doi.org/10.1016/j.intell.2017.01.012

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A

quantitative review. *Psychological Bulletin, 137,* 856–879. http://dx.doi.org/10.1037/a0023678

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., . . . Salum, G. A. (2017). A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology, 126,* 137–148. http://dx.doi.org/10.1037/abn0000205

McElroy, E., Belsky, J., Carragher, N., Fearon, P., & Patalay, P. (2018). Developmental stability of general and specific factors of psychopathology from early childhood to adolescence: Dynamic mutualism or p-differentiation? *Journal of Child Psychology and Psychiatry, 59,* 667–675. http://dx.doi.org/10.1111/jcpp.12849

Michelini, G., Barch, D. M., Tian, Y., Watson, D., Klein, D. N., & Kotov, R. (2019). Delineating and validating higher-order dimensions of psychopathology in the Adolescent Brain Cognitive Development (ABCD) study. *Translational Psychiatry, 9,* 261. http://dx.doi.org/10.1038/s41398-019-0593-4

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (Eighth ed.). Los Angeles, CA: Author.

Neumann, A., Pappa, I., Lahey, B. B., Verhulst, F. C., Medina-Gomez, C., Jaddoe, V. W., . . . Tiemeier, H. (2016). SNP heritability of a general psychopathology factor in children. *Journal of the American Academy of Child & Adolescent Psychiatry, 55,* 1038–1045. http://dx.doi.org/10.1016/j.jaac.2016.09.498

Olino, T. M., Bufferd, S. J., Dougherty, L. R., Dyson, M. W., Carlson, G. A., & Klein, D. N. (2018). The development of latent dimensions of psychopathology across early childhood: Stability of dimensions and moderators of change. *Journal of Abnormal Child Psychology, 46,* 1373–1383. http://dx.doi.org/10.1007/s10802-018-0398-6

Pettersson, E., Lahey, B. B., Lundström, S., Larsson, H., & Lichtenstein, P. (2018). Criterion validity and utility of the general factor of psychopathology in childhood: Predictive associations with independently measured severe adverse mental health outcomes in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 57,* 372–383.

Quay, H. C. (1986). Classification. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (3rd ed., pp. 1–42). New York, NY: Wiley.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47,* 667–696. http://dx.doi.org/10.1080/00273171.2012.715555

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21,* 137–150. http://dx.doi.org/10.1037/met0000045

Sallis, H., Szekely, E., Neumann, A., Jolicoeur-Martineau, A., van IJzendoorn, M., Hillegers, M., . . . Evans, J. (2019). General psychopathology, internalising and externalising in children and functional outcomes in late adolescence. *Journal of Child Psychology and Psychiatry, 60,* 1183–1190. http://dx.doi.org/10.1111/jcpp.13067

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31,* 1428–1441. http://dx.doi.org/10.1037/pas0000623

Settles, R. E., Fischer, S., Cyders, M. A., Combs, J. L., Gunn, R. L., & Smith, G. T. (2012). Negative urgency: A personality predictor of externalizing behavior characterized by neuroticism, low conscientiousness, and disagreeableness. *Journal of Abnormal Psychology, 121,* 160–172. http://dx.doi.org/10.1037/a0024948

Shields, A. N., Reardon, K. W., Brandes, C. M., & Tackett, J. L. (2019). The p factor in children: Relationships with executive functions and effortful control. *Journal of Research in Personality, 82,* 103853. http://dx.doi.org/10.1016/j.jrp.2019.103853

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society Series B, Statistical Methodology, 13,* 238–241.

Snyder, H. R., Young, J. F., & Hankin, B. L. (2017). Strong homotypic continuity in common psychopathology-, internalizing-, and externalizing-specific factors over time in adolescents. *Clinical Psychological Science, 5,* 98–110. http://dx.doi.org/10.1177/2167702616651076

Sunderland, M., Forbes, M. K., Mewton, L., Baillie, A. J., Carragher, N., Lynch, S., . . . Slade, T. (2019). The structure of psychopathology and association with poor sleep, self-harm, suicidality, risky sexual behaviour, and low self-esteem in a population sample of adolescents. *PsyArXiv.* http://dx.doi.org/10.31234/osf.io/5f4vj

Taylor, E. M. (1959). *The appraisal of children with cerebral deficits.* Cambridge, MA: Harvard University Press.

Tulsky, D. S., Carlozzi, N. E., Chevalier, N., Espy, K. A., Beaumont, J. L., & Mungas, D. (2013). V. NIH Toolbox Cognition Battery (CB): Measuring working memory. *Monographs of the Society for Research in Child Development, 78,* 70–87. http://dx.doi.org/10.1111/mono.12035

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41,* 321–327. http://dx.doi.org/10.1007/BF02293557

Volkow, N. D., Koob, G. F., Croyle, R. T., Bianchi, D. W., Gordon, J. A., Koroshetz, W. J., . . . Weiss, S. R. B. (2018). The conception of the ABCD study: From substance use to a broad NIH collaboration. *Developmental Cognitive Neuroscience, 32,* 4–7. http://dx.doi.org/10.1016/j.dcn.2017.10.002

Watts, A. L., Poore, H. E., & Waldman, I. D. (2019). Riskier tests of the validity of the bifactor model of psychopathology. *Clinical Psychological Science, 7,* 1285–1303.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children* (5th ed.). Bloomington, MN: Pearson.

Zapolski, T. C. B., Stairs, A. M., Settles, R. F., Combs, J. L., & Smith, G. T. (2010). The measurement of dispositions to rash action in children. *Assessment, 17,* 116–125. http://dx.doi.org/10.1177/1073191109351372

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1,* 297–301. http://dx.doi.org/10.1038/nprot.2006.46

Correction to Moore et al. (2020)