Construct Validity of the BASC-3 Teacher Rating Scales: Independent Hierarchical Exploratory Factor Analyses with the Normative Sample

Gary L. Canivez[1], Nathaniel P. von der Embse[2], and Ryan J. McGill[3]

1Department of Psychology, Eastern Illinois University

[2]University of South Florida

[3]Department of School Psychology and Counselor Education, William & Mary

Author Note

Gary L. Canivez https://orcid.org/0000-0002-5347-6534

Nathaniel P. von der Embse https://orcid.org/0000-0003-2545-7691

Ryan J. McGill https://orcid.org/0000-0002-5138-0694

We have no conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Gary L. Canivez, Department of Psychology, Eastern Illinois University, Charleston, IL 61920. E-mail: glcanivez@eiu.edu

This research was partially supported by a 2018 Summer Research/Creative Activity Award to the first author. Preliminary results were presented at the 2019 meeting of the National Association of School Psychologists, Atlanta, GA.

Abstract

The Behavior Assessment System for Children-Third Edition (BASC–3) is the most recent edition and the Teacher Rating Scales (TRS) was reported to be the most frequently used test in school psychology practice. Despite its popularity, there is a lack of independent empirical research regarding psychometric properties. The BASC–3 *Manual*, while quite detailed in many respects, lacks important details in reporting TRS item and scale level factor analyses limiting confidence in construct validity based on internal structure. The present study examined the latent factor structure of the BASC–3 TRS Preschool, Child, and Adolescent Clinical and Adaptive scales using best practices in exploratory factor analysis (EFA). EFA was conducted with the Clinical *and* Adaptive scales jointly, and with the Clinical scales separately, to aid interpretive clarity. Results indicated theoretically consistent alignment of the BASC–3 TRS Clinical scales to their specified factors (Externalizing, Internalizing, and School Problems) and an additional factor (Social Disengagement) was identified suggesting a possible new latent construct for a composite scale score containing the Withdrawal and Atypicality scales. Variance partitioning applied to second-order EFA and model-based validity statistics, however, indicated that the composite scales (Externalizing, Internalizing, School Problems, and Social Disengagement) appear to lack sufficient unique variance for confident clinical interpretation in isolation.

Keywords: BASC–3, teacher rating scales, higher-order exploratory analysis, factor structure, Schmid-Leiman

Impact Statement

This study reports the first independent evaluation of the dimensions measured by the Behavior Assessment System for Children-Third Edition (BASC–3) Teacher Rating Scales (TRS). Results partially supported the interpretive structure and a possible new dimension (Social Disengagement) was identified.

Construct Validity of the BASC–3 Teacher Rating Scales: Independent Hierarchical Exploratory Factor Analyses of the Normative Sample

The Behavior Assessment System for Children-Third Edition (BASC–3; Reynolds & Kamphaus, 2015a) is the most recent edition and includes the Behavioral and Emotional Screening System (BESS), Teacher Rating Scales (TRS), Parent Rating Scales (PRS), Self-Report of Personality (SRP), Structured Developmental History (SDH), Student Observation System (SOS), and other features (Reynolds & Kamphaus, 2015b). According to a recent survey of school psychologist test usage (Benson et al., 2019) the BASC–3 TRS was the most frequently used test in school psychology practice. Despite the popularity and use of the BASC–3 TRS, there has been a lack of independent, empirical, and peer-reviewed factor analytic research that would support the assertions and preliminary psychometric results presented within the BASC–3 *Manual*.

Construct validity is a crucial element for judging the adequacy of test interpretation. One of the most important elements of construct validity is that based on a test's internal structure (Messick, 1995) because it is from that structure that scales and composite scores are derived and used for interpretation (American Educational Research Association [AERA] et al., 2014). Kane (2013) advanced an argument for the need to include evidence of a test's *interpretation* and *use.* Evidence for test interpretation may come from construct validity (e.g., factor structure, measurement invariance, convergent and discriminant validity) whereas evidence for score use would support the *decisional* inferences of the derived scores (Kane, 2013). In addition to test structure, the measurement contribution of provided scores from multidimensional tests requires assessment *within* the test using indices such as Omega-hierarchical (ω_H) and Omega-hierarchical subscale (ω_{HS}) (Reise, 2012); construct replicability (*H*) (Hancock & Mueller, 2001), and the factor determinacy index (FDI) (Rodriguez et al., 2016a, b). Test structure and related scores must also be evaluated with criteria *external* to the test (e.g., predictive validity, incremental validity [Hunsley, 2003], and diagnostic utility [Kessel & Zimmerman, 1993]).

The BASC–3 TRS offers a panoply of scores, some of which emerge from item content created to measure various scales and selected through item factor analyses and it is these scales that are the focus of the present study. Beaujean and Benson (2019), in the context of intelligence test development and interpretation of the many available scores, noted “the problem with offering such a bevy of scores to interpret is that there is no single psychometric or attribute theory that can support all their interpretations” (p. 126). Hence, there is a need for tests to be developed based first on well-defined and described attributes and the theory(ies) from which they emanate. Reynolds and Kamphaus (2015b) noted that BASC–3 development included scales that were clearly conceptualized, and balanced theory and empirical findings. BASC-3 development considered symptoms codified in the International Classification of Diseases (ICD-10), the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013), the Individuals with Disabilities Educational Improvement Act (IDEA, 2004), the Americans with Disabilities Act (ADA, 1990), and Section 504 (Rehabilitation Act, 1973) so the BASC–3 could be interpreted in different environments. Factor analyses of scales resulted in factor composite scales related to dimensions frequently observed in the child psychopathology literature (Externalizing and Internalizing). Other scores and indexes have also been created (Content scales [e.g., Bullying, Executive Functioning, Resiliency] and Indexes [e.g., ADHD Probability Index, Autism Probability Index, Attentional Control Index]) for various purposes but were not factorially derived and included items from different scales and thus not specifically examined in the present study. It was noted in the BASC–3 *Manual* that the Content scales were “initially developed based on theory and expert review” Reynolds and Kamphaus (2015, p. 4). Thus, it appears that a variety of theories have been applied though it is not clear which.

BASC TRS Factor Structure Research

The first Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) and its revision with new norms, the Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004) both reported similar structural validity methods and results in their respective *Manuals* and the BASC–3 *Manual* reported these earlier versions as starting points for item revision and introduction of new items to improve TRS measurement. Neither the BASC-2 *Manual* nor the BASC–3 *Manual* included any TRS structural validity studies for earlier versions of the BASC in the independent peer-reviewed literature and our review of the literature was unable to locate any.

Item Level Analyses. Despite the preliminary psychometric support of reliability and validity of the BASC–3 TRS as reported in the *Manual*, there are four primary shortcomings. First, psychometric methods and properties for *item* level analyses were inadequately reported, thus independent evaluation of item structure is not possible. Univariate and multivariate descriptive statistics, including skewness and kurtosis, were not provided and if not normally distributed special analyses such as robust model estimation would be needed. Second, the exclusive use of "Covariance Structure Analysis (CSA; also known as confirmatory factor analysis)" (Reynolds & Kamphaus, 2015b, p. 88) was reportedly used in scale development, including use of modification indexes, with the TRS item development sample that included combined normative *and* clinical samples (Preschool [n = 800], Child [n not specifically reported but is somewhere between 705 and 1,330 as only range across the Child sample was reported], and Adolescent [n = 956]). While confirmatory factor analyses (CFA) may be used, full disclosures of methods are necessary to assess model adequacy and results (Applebaum et al., 2018; Schreiber et al., 2006). The *method* of CFA, and in turn the related assumptions, were not provided including no report of what estimator was used, how the scales were identified, or whether item data were treated as categorical or ordinal. Third, because BASC–3 TRS items are categorical and follow an ordinal (graded-response) scale, item data should be treated as such in CFA (Li, 2016).

Fourth, CFA fit statistics for item level analyses were absent making it impossible to judge model adequacy. With items assigned to separate scales, CFA might not adequately assess the extent to which items had significant associations with other scales (i.e., cross-loading) and could influence clinical interpretation. Also, if polychoric correlations were not used, item loadings and cross-loadings may differ from those reported. The BASC-3 *Manual* noted the use of AMOS 6 for CSA/CFA analyses, but AMOS 6 does not provide polychoric/polyserial correlations.[1] This is particularly problematic given the widespread use of BASC–3 TRS scales to inform decision-making and treatment selection. Table 7.1 in the BASC–3 *Manual* presents TRS item standardized factor loadings in final analyses and similar associations of items to their assigned scale appeared generally supportive. However, while CFA can be used, item level exploratory factor analyses (EFA) would help identify problematic indicators and the optimal number of factors to extract. Although somewhat consistent factor structures were reported in previous BASC TRS versions, there is a need for independent evaluation of the reported factor structure to support interpretation given the lack of peer-reviewed evidence replicating findings across earlier versions of the test.

Scale Level Analyses. There are several additional concerns regarding reported scale level analyses, including inconsistencies with widely utilized thresholds for acceptable model fit. At the scale level, CFA was again used to assess the latent factor structure of the BASC–3 TRS Clinical and Adaptive scales using the item-development sample containing combined clinical and normative samples that were also used in item level CFA. Like item level CFA, scale level

[1] https://www.ibm.com/support/pages/does-amos-use-polychoric-correlations

CFA lacked sufficient disclosure of CFA methods. It is unknown if scales reflected univariate and multivariate normality, what estimator was used, or how the scales were identified. Table 9.16 in the BASC–3 *Manual* presents some TRS CFA fit statistics that were reported as "moderate" (Reynolds & Kamphaus, 2015b, p. 139) in overall fit; however, the fit statistics for the TRS Preschool ($\chi^2(41)$ = 1,383.9, CFI = .76, RMSEA = .20), Child ($\chi^2(83)$ = 3,056.2, CFI = .81, RMSEA = .18), and Adolescent ($\chi^2(83)$ = 2,632.4, CFI = .82, RMSEA = .18) were not acceptable or good compared to consensus standards (e.g., Brown, 2014).

Need for Independent Evaluation of Reported Factor Structures

Using the factor loadings for the final TRS models reported in Table 9.17 in the BASC–3 *Manual* (Reynolds & Kamphaus, 2015b, p. 140), standardized measurement models were constructed and are illustrated in Figures 1-3. For the TRS Preschool, the measurement model reflects simple structure with each BASC–3 scale loading on a single latent factor (Externalizing, Internalizing, Adaptive Skills). For both the TRS Child and TRS Adolescent, measurement models did not reflect simple structure with the Adaptability scale having primary loading on Adaptive Skills but also a secondary (and negative) loading on Internalizing. There was no discussion in the BASC–3 *Manual* as to why this additional path was added and if it was an *a priori* decision or a *post hoc* decision possibly suggested by modification indexes. What is also illustrated in Figures 1-3 is that the latent factors contain high factor correlations that may indicate hierarchical structure that should also be examined for full understanding of the latent structure (Canivez, 2016; Gorsuch, 1983). Alternate or rival TRS measurement models were also not presented in the BASC–3 *Manual*.

The BASC–3 *Manual* also presents EFA results using principal-axis extraction with the TRS Preschool, Child, and Adolescent scales. While it was reported that results were similar for oblique and orthogonal rotations, only the orthogonal rotation (Varimax) results were presented.

This is perplexing given the large factor correlations presented for the TRS CFAs reported in the BASC–3 *Manual*, as well as Composite (Factor) score correlations presented in Tables 9.13, 9.14, and 9.15, clearly showing significant factor correlations. Thus, Varimax rotation is likely not appropriate (Gorsuch, 1983). Without presenting the oblique factor pattern and structure coefficients (Thompson, 2004) it is not possible to independently verify how similar the two rotation methods were. As Gorsuch (1983) noted, "implicit in all oblique rotations are higher-order factors. It is recommended that these be extracted and examined so that the investigator may gain the fullest possible understanding of the data" (p. 255). Relatedly, a hierarchical structure requires partitioning variance to higher- and lower-ordered dimensions to understand adequacy of measurement of multidimensional constructs, but this also was not provided. There appears to be no variance estimates provided for the TRS in either CFA or EFA analyses in the BASC–3 *Manual,* which limits test user ability to judge the adequacy of various scores.

One final note regarding factor analyses of the BASC–3 TRS relates to the separation of *item* level factor analyses and *scale* level factor analyses. In measures of psychopathology (unlike intelligence tests) *items* would be considered the measured variables or indicators (measures of various behaviors or emotions) and items measuring a particular latent trait or characteristic would have higher correlations (convergent validity) and in factor analysis represent a factor. In a scale such as the BASC–3 TRS, several factors emerge from item clusters sharing variance and thus the emergence of multiple first-order factors (scales). Various theoretically related scales (i.e., Hyperactivity, Aggression, and Conduct Problems; Anxiety, Depression, and Somatization) would be expected to have higher correlations with each other (convergent validity) and lower correlations with other scales (discriminant validity), and second-order factor analysis might identify higher-order factors (i.e., externalizing problems and internalizing problems) such that the correlated scales could be combined into a composite score representing a broader (second-order) domain. Thus, it would be informative if EFA *began* with the BASC–3 TRS *items* to extract the first-order factors (scales) and using oblique rotation, identify the first-order factor correlations that would be used in second-order EFA to identify the higher-order factors that would represent the composites. Where second-order factors are correlated it is possible that a third-order EFA with oblique rotation could allow for examination of even higher-order dimensions (Wolff & Preising, 2005). Because the BASC–3 TRS items are rated on a four-point ordinal (graded-response) rating scale, polychoric correlations might be a more proper item level correlation method for first-order EFA, particularly in the presence of non-normally distributed data (Sellbom & Tellegen, 2019). CFA procedures would be similar, specifying BASC–3 TRS items as categorical or ordinal indicators and various models beyond oblique (correlated) factors (higher-order or bifactor) could be specified and compared. Therefore, despite the widespread use of the BASC–3 TRS and preliminary support in the BASC–3 *Manual*, there is a need for independent research to evaluate scale *interpretation*. Unfortunately, item level raw data were not available for independent analyses in the present study.

Present Study

Beaujean (2015) opined a revised test should be treated like a new test as it cannot be assumed that scores from the revision would be directly comparable to the previous version without supporting evidence. Given the limitations of the reported structural validity evidence for the TRS within the BASC–3 *Manual,* including the reported CFA and EFA methods employed, and the apparent lack of independent factor analytic research on previous BASC editions, it was necessary to independently examine the latent factor structure of the BASC–3 TRS scales. The primary aims of the present study included 1) a determination of the adequacy of BASC–3 TRS factor structure using best practices in EFA (Watkins, 2018), 2) an estimation of the portions of variance attributed to the various factors and scales, and 3) an evaluation of reported alpha coefficients concerning multidimensionality or other violations of alpha assumptions (Watkins, 2017). Model-based validity estimates (ω_H, ω_{HS}, *H,* FDI) were used to provide estimates of *unique* contributions of variance necessary to judge adequacy of factorially derived scores.

Method

Participants

Participants in the present study were from the BASC–3 TRS Preschool ($N = 500$), Child ($N = 600$), and Adolescent ($N = 600$) general standardization norm samples and demographic characteristics are presented and described in the BASC–3 *Manual.* Standardization samples were reportedly demographically representative of the U.S. population across variables of parent education level (a proxy for SES), race/ethnicity (African American, Asian, Hispanic, White, Other), and geographic region based on 2013 U.S. census data.

Instrument

The BASC–3 provides multidimensional behavioral assessment of both adaptive skills and clinical problems. Similar scales are present across the Preschool, Child, and Adolescent versions but some scales are unique based on developmental level. The BASC–3 TRS Preschool includes Hyperactivity and Aggression scales and their Externalizing Problems Composite; Anxiety, Depression, and Somatization scales and their Internalizing Problems Composite; Attention Problems, Atypicality, and Withdrawal scales; and Adaptability, Social Skills, and Functional Communication scales and their Adaptive Skills Composite. The BASC–3 TRS Child and BASC–3 TRS Adolescent versions include Hyperactivity, Aggression, and Conduct Problems scales and their Externalizing Problems Composite; Anxiety, Depression, and Somatization scales and their Internalizing Problems Composite; Attention Problems and

Learning Problems and their School Problems Composite; Atypicality and Withdrawal scales; and Adaptability, Social Skills, Leadership, Study Skills, and Functional Communication scales and their Adaptive Skills Composite. All three BASC–3 TRS versions include a Behavioral Symptoms Index that is an atheoretical composite score that includes various scales from Externalizing, Internalizing, and School Problems domains and scales without a factor based composite score (Atypicality, Withdrawal, and Attention Problems [Preschool]).[2]

Reliability estimates for the BASC–3 TRS Preschool, Child, and Adolescent forms are provided BASC–3 *Manual* and indicated generally supportive results for score consistency. Validity estimates (relationships with other measures and clinical group characteristics) also provided some preliminary support for scale interpretation. As previously noted, evidence presented in the BASC–3 *Manual* for the internal structures of the BASC–3 TRS versions was less than satisfactory.

Procedure

Adaptive and Clinical scales correlation matrices for the TRS samples were obtained from BASC–3 *Manual* Tables 10.15, 10.16, and 10.17, for these analyses.[3] While the reported CFA and EFA of TRS in the BASC–3 *Manual* were conducted with the combined normative *and* clinical samples, correlation matrices published in the BASC–3 *Manual* are separately presented for the *normative* samples (below diagonal) or *clinical* samples (above diagonal). While the

---

[2] The BASC–3 also provides theoretical or syndrome-oriented "Content Scales" (e.g., Negative Emotionality, Bullying) that include item content from different BASC–3 scales as well as items uniquely created for the Content Scale. They were not factorially derived and like analyses reported in the BASC–3 *Manual*, not included in the present analyses.

[3] Standardization sample item raw data and demographic data for the BASC–3 TRS Preschool, Child, and Adolescent standardization samples were requested from NCS Pearson, Inc., by the first author after encouragement by the second-author of the BASC-3 (R. Kamphaus, personal communication, August 9, 2015), in order to conduct *item* level factor analyses but access to item and scale raw data was denied by the test publisher. Because the BASC–3 TRS is a measure of child psychopathology, *items* would be considered the measured variables to be the starting point in EFA and CFA. Also, BASC–3 TRS items are rated on a four-point ordinal (graded-response) rating scale (*Never, Sometimes, Often, Very Often*) and likely deviate from normal distribution, thus polychoric item correlations should be used rather than Pearson product-moment correlations in item level factor analyses.

BASC–3 *Manual* does not specify which correlations are above or below the diagonal, a representative from Pearson Customer Service verified this configuration with the psychometrics department via email inquiry by the first author (personal communication, April 23, 2018). Thus, present analyses were with the normative sample correlations (below diagonal) reported in Tables 10.15, 10.16, and 10.17).

Data Analyses

Exploratory Factor Analyses. Best practices in EFA as outlined by Watkins (2018) were used. Principal axis factoring (Fabrigar & Wegener, 2012) was used to analyze reliable common variance from the BASC–3 TRS correlation matrices using SPSS 24.0 for Macintosh. Correlation matrices were evaluated with Bartlett's Test of Sphericity (Bartlett, 1954) to determine if matrices were not random and the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (Kaiser, 1974) with a .60 minimum standard to determine matrix factorability. Multiple criteria were examined and considered for suggesting the number of latent factors to retain (Gorsuch, 1983) and included eigenvalues > 1 (Guttman, 1954), the visual scree test (Cattell, 1966), standard error of scree (SE_{Scree}; Zoski & Jurs, 1996), Horn's parallel analysis (HPA; Horn, 1965) including Glorfeld's (1995) modification, and minimum average partials (MAP, Velicer, 1976). The scree test is a subjective criterion where the optimum number of factors to retain is visually determined and the SE_{Scree} was used as programmed by Watkins (2007) as it was reported to be the most accurate objective scree method (Nasser et al., 2002). HPA has been shown to be one of the most accurate *a priori* empirical criteria with simulation studies suggesting scree is sometimes useful (Velicer et al., 2000). HPA and Glorfeld's modification (95% CI) were included as typically more accurate and reducing overfactoring (Frazier & Youngstrom, 2007). However, Crawford et al. (2010) suggested that HPA tends to suggest fewer factors in the presence of a strong general factor. HPA indicated potentially meaningful factors when sample data eigenvalues exceeded eigenvalues produced by random data containing the same number of participants and factors. The *Monte Carlo PCA for Parallel Analysis* computer program (Watkins, 2000) with 100 replications was used to provide stable eigenvalue estimates for random data for HPA. MAP was conducted using the O'Connor (2000) SPSS syntax.

Promax rotation (k = 4 [to maximize hyperplane count]; Gorsuch, 1983) was used following extraction to examine correlated factors and viable factors required a minimum of two scales with salient factor pattern coefficients (≥ .40). It was also preferable to achieve simple structure (i.e., no scale cross-loadings; Thurstone, 1947). In four instances communality estimates exceeded 1 in iterations (indicative of a Heywood case) so extraction iterations were limited to 2 as per Gorsuch (2003) for model estimation. Scales with factor pattern coefficients between .30 and .39 were considered "aligned" with an extracted factor for descriptive purposes when failing to achieve saliency. First unrotated factor structure coefficients were examined to assess BASC–3 TRS scale general factor saturation and Kaufman's (1994) criteria (≥ .70 = good, .50–.69 = fair, < .50 = poor) were applied.

Second-Order Exploratory Factor Analyses.

Second-order EFA was conducted using promax rotated first-order factor correlations and results then transformed using the Schmid and Leiman (SL; 1957) orthogonalization procedure. Using the obliquely rotated first-order factor pattern coefficients and second-order EFA solutions, the SL procedure based on SPSS syntax code (Wolff & Preising, 2005) apportioned common variance first to the second-order factor and the residual common variance was then apportioned to the first-order (group) factors to better understand sources of variability within the various BASC–3 TRS scales. The SL procedure is a reparameterization of the higher-order model (Reise, 2012) and "not only preserves the desired interpretation characteristics of the oblique solution, but also discloses the hierarchical structuring of the variables" (Schmid & Leiman, 1957, p. 53).

Model Based Validity Analyses. Omega–hierarchical (ω_H) and omega–hierarchical subscale (ω_{HS}) coefficients (Reise, 2012) were estimated as model–based validity coefficients of the latent factors (Gignac & Watkins, 2013). Chen et al. (2012) noted that "for multidimensional constructs, the alpha coefficient is complexly determined, and McDonald's omega–hierarchical (ω_H; 1999) provides a better estimate for the composite score and thus should be used" (p. 228). ω_H is the estimate of interpretive relevance for a hierarchical general factor independent of the variance of group factors, while ω_{HS} is the estimate of interpretive relevance of a group factor with all other group *and* general factors removed (Reise, 2012). Omega estimates (ω_H and ω_{HS}) may be obtained from EFA SL solutions and were produced using the *Omega* program (Watkins, 2013), which is based on the tutorial by Brunner et al. (2012). Omega coefficients should exceed .50, but .75 is preferred (Reise, Bonifay, et al., 2013). Omega coefficients were supplemented by the *H* coefficient (Hancock & Mueller, 2001) that is a construct replicability coefficient and the correlation between a factor and an optimally weighted composite score. *H* indexes how well the latent factor is represented by the indicators and the recommended minimum criterion value of .70 (Rodriguez et al., 2016a, b) was used. The factor determinacy index (FDI) was also used to determine how well the underlying factor was estimated by the factor scores and a criterion value ≥ .90 (Rodriguez et al., 2016a, b) was used. Scale specificity estimates were calculated by subtracting general and group factor variance (communality) from the median TRS scale alpha coefficient obtained from the BASC–3 *Manual*. Kaufman and Lichtenberger (2005) criteria for scale specificity interpretation were applied where specificity ≥ .25 *and* > error variance was *ample,* while specificity < .25 *but* > error variance might be *adequate*. Finally, unidimensionality was considered with explained common variance (ECV) as a proportion of common variance explained by the target construct and the percentage of uncontaminated correlations (PUC) as an indication of bias that might "result from forcing multidimensional data into a unidimensional model" (Watkins, 2013, no page). ECV values of .70 to .80 and PUC ≥ .80 might indicate essential or sufficient unidimensionality (Rodriguez et al., 2016a) and noted "When ECV is > .70 and PUC > .70 relative bias will be slight and the common variance can be regarded as essentially unidimensional" (p. 232). All model-based coefficients were produced by the *Omega* program (Watkins, 2013).

Results

Due to a large number of tables and figures presenting results from the many analyses, only those of primary importance are presented here. Additional tables and figures from analyses less important are presented in an Appendix available as an online supplement and designated with an A (i.e., Table A1, Figure A1, etc.) for complete reporting.

BASC–3 TRS Preschool

Clinical and Adaptive Scales EFA

Bartlett's Test of Sphericity, $\chi^2(55) = 3{,}389.89$, $p < .0001$; indicated the correlation matrix was not random and the KMO Measure of Sampling Adequacy of .841 far exceeded the minimum standard for factorability. Initial communality estimates ranged from .257 to .752 (*Mdn* = .600). Table 1 presents the number of factors suggested by various criteria and indicated between three and five factors might be extracted. EFA began with extraction of 5 factors and iteratively reduced factor extraction by one to examine resulting structures.

Clinical and Adaptive scales EFA with 5 through 2 extractions produced inadequate results producing single scale associations (5 factors and 4 factors [Table A1]), salient scale factor pattern coefficients on multiple factors (cross-loading; 5, 4, and 3 factors [Table A1 and A2]), and scales with no salient pattern coefficients on any factor (3 factors and 2 factors).

Extraction of only two factors produced merging of scales from Internalizing and Externalizing factors into a general Problem Behaviors factor. These results were inconsistent with the structure purported in the *Manual* and may be the result of conducting EFA with bipolar oriented scales in the same analyses. Future research with item and scale level raw data will no doubt be necessary to completely replicate the reported factor structure. A final observation was that all BASC–3 TRS Preschool scales, except Somatization, showed fair to good first unrotated factor structure coefficients using Kaufman's (1994) criteria, indicating general factor saturation.

Clinical Scales EFA

Bartlett's Test of Sphericity, $\chi^2(28) = 2{,}168.9$, $p < .0001$; indicated the correlation matrix was not random and the KMO Measure of Sampling Adequacy of .796 far exceeded the minimum standard for factorability. Initial communality estimates ranged from .238 to .736 (*Mdn* = .534). Factor extraction criteria (see Table 1) indicated that the BASC–3 TRS Preschool Clinical scales might produce two to four factors. Thus, extraction began with four factors and iteratively reduced extractions by one to examine structures.

Extraction of four factors (Table A3) was inadequate with the fourth factor defined by Aggression and Depression scales and Aggression also had salient factor pattern coefficients on two factors (cross-loading). This may be a result of overextraction (Gorsuch, 1983). The most plausible solution was the three-factor extraction (see Table 2) with Hyperactivity, Aggression, and Attention Problems saliently loading the Externalizing factor; Anxiety, Depression, and Somatization saliently loading the Internalizing factor; and Withdrawal and Atypicality saliently loading a new factor tentatively named Social Disengagement. The three-factor model satisfied factor selection criteria including simple structure. An alternative two factor model (see Table 3) that is proposed for the BASC–3 TRS Preschool scales also achieved simple structure and included Externalizing and Internalizing factors with Withdrawal and Atypicality joining the

Internalizing factor. However, inspection of the factor structure coefficients for Withdrawal and Atypicality on the Social Disengagement factor (see Table 2) were higher than they were on the Internalizing factor (see Table 3). As observed in analyses with Clinical and Adaptive scales, all BASC–3 TRS Preschool Clinical scales, except Somatization, showed fair to good first unrotated factor structure coefficients using Kaufman's (1994) criteria, indicating general factor saturation.

TRS-Preschool Clinical Scales Second-Order EFA With SL Transformation

Three Group Factors. Due to statistically significant and moderate to large factor correlations for the three-factor model (Table 2), second-order EFA of these factor correlations was performed and results subjected to the SL orthogonalization procedure. Table 4 presents the resulting variance decomposition and metrics to assess first- and second-order factor interpretability. The second-order general factor accounted for 36.0% of total variance and 54.5% of common variance, while the three first-order factors (Externalizing, Internalizing, and Social Disengagement) accounted for considerably less unique variance (5.8% to 15.1% of total variance, 8.8% to 22.8% common variance). Omega coefficients were estimated based on the decomposed variance estimates in Table 4 and the ω_H coefficient for the general factor (.716) was high and sufficient for confident scale interpretation of a unit-weighted score based on all BASC–3 TRS Preschool Clinical scales, but the ω_{HS} coefficients for the Externalizing, Internalizing, and Social Disengagement factors (.244-.449) did not reach the minimum threshold for confident interpretation of unit-weighted scores based on specified indicators. The *H* indexes indicated an optimally weighted composite score for a general factor would account for 82.9% of variance, but the three first-order factors were not well defined by their optimally weighted indicators (*H*s < .70). ECV and PUC values for the general factor indicated that the general factor was not sufficiently unidimensional. Another assessment of possible interpretation rests with the scale specificity estimates in Table 4 that illustrated that the Attention Problems and Somatization scales contained ample unique variance and the Aggression and Atypicality scales contained adequate unique variance for possible separate interpretation. Hyperactivity, Anxiety, Depression, and Withdrawal scales did not have adequate specificity for separate interpretation.

Two Group Factors. Due to the statistically significant and large factor correlation for the two-factor model (Table 3), second-order EFA with this factor correlation was performed and the SL orthogonalization procedure applied to results. Table 5 presents the resulting variance decomposition and metrics to assess first- and second-order factor interpretability. The second-order general factor accounted for 34.0% of total variance and 57.7% of common variance while the two first-order factors (Externalizing, Internalizing) accounted for considerably less unique variance (11.7% and 13.2% of total variance, 19.8% and 22.5% common variance). Omega coefficients were estimated based on the decomposed variance estimates in Table 5 and the ω_H coefficient for the general factor (.668) met the minimum criterion for scale interpretation of a unit-weighted score based on all BASC–3 TRS Preschool Clinical scales but slightly less than the preferred value (.70). ω_{HS} coefficients for the Externalizing (.387) and Internalizing (.299) factors did not reach the minimum threshold for confident interpretation of unit-weighted scores based on specified indicators. *H* indices indicated an optimally weighted composite score for a general factor would account for 81.8% of variance, but the two first-order factors were not well defined by their optimally weighted indicators (*H*s < .70). ECV and PUC estimates indicated that the general factor was not sufficiently unidimensional. Another assessment of potential interpretation reflected by the scale specificity estimates in Table 5 illustrated that the Attention Problems, Withdrawal, Somatization, and Atypicality scales contained ample unique variance while the Aggression scale contained adequate unique variance for possible separate interpretation. Hyperactivity, Anxiety, and Depression scales did not have adequate specificity for separate interpretation. Reducing the number of extracted factors to two resulted in higher scale specificity estimates.

BASC–3 TRS Child

Clinical and Adaptive Scales EFA

Bartlett's Test of Sphericity, $\chi^2(105) = 7{,}494.38$, $p < .0001$; indicated the correlation matrix was not random and the KMO Measure of Sampling Adequacy of .896 far exceeded the minimum standard for factorability. Initial communality estimates ranged from .223 to .838 (*Mdn* = .697). Various criteria and suggested extraction of between two and five factors (Table 1). EFA began with extraction of 5 factors and iteratively reduced extraction by one to examine resulting structures.

Extracting five factors (see Table A4) produced suboptimal results with Learning Problems and Attention Problems having salient positive factor pattern coefficients and Study Skills, Leadership, and Functional Communication having salient *negative* factor pattern coefficients on the same factor; and Leadership and Withdrawal had salient factor pattern coefficients on two factors. Extraction of four factors (see Table A5) was also suboptimal as Atypicality had no salient factor pattern coefficients on any factor and Attention Problems, Functional Communication, and Leadership had salient factor pattern coefficients on two factors. The most optimal solution was the three-factor extraction (Table A6) that produced simple structure, however, Factor 1 included salient factor pattern coefficients from all Adaptive scales, but also included salient *negative* factor pattern coefficients from Learning Problems, Attention Problems, Withdrawal, and Atypicality. Factor 2 was an Externalizing factor with salient factor pattern coefficients with Aggression, Conduct Problems, Hyperactivity; and Factor 3 was an Internalizing factor with salient factor pattern coefficients with Anxiety, Depression, Somatization). The least optimal solution was a two-factor model with Somatization and Anxiety having no salient factor pattern coefficients on either factor. Factor 1 included all Adaptive scales that had salient positive factor pattern coefficients and Learning Problems, Attention Problems, Withdrawal, and Atypicality had salient *negative* factor pattern coefficients. Like the BASC–3 TRS Preschool, problems observed in BASC–3 TRS Child EFA including both Clinical and Adaptive scales may be the result of including both positive and negative scales in the same analyses. A final observation was that all BASC–3 TRS Child Clinical and Adaptive scales, except Anxiety and Somatization, showed fair to good first unrotated factor structure coefficients using Kaufman's (1994) criteria, indicating saturation of a general factor.

Clinical Scales EFA

Bartlett's Test of Sphericity, $\chi^2(45) = 3{,}716.47$, $p < .0001$; indicated the correlation matrix was not random and the KMO Measure of Sampling Adequacy of .816 far exceeded the minimum standard for factorability. Initial communality estimates ranged from .206 to .773 ($Mdn = .623$). Factor extraction criteria (see Table 1) indicated that the BASC–3 TRS Child Clinical scales might produce two to four factors. Thus, extraction began with four factors and iteratively reduced extractions by one to examine structures.

Extraction of four factors (Table 6) was the most plausible solution and produced the desired simple structure. Aggression, Conduct Problems, and Hyperactivity had salient factor pattern coefficients on the Externalizing factor; Attention Problems and Learning Problems had salient factor pattern coefficients on the School Problems factor; Withdrawal and Atypicality had salient factor pattern coefficients on the Social Disengagement factor; and Anxiety, Somatization, and Depression had salient factor pattern coefficients on the Internalizing factor. Extraction of three factors (see Table A7) consistent with the purported BASC–3 TRS Child structure was inadequate as Withdrawal had no salient factor pattern coefficients on any factor and Atypicality had a salient factor pattern coefficient on the School Problems factor. Extraction of two factors produced somewhat plausible structure with Hyperactivity, Conduct Problems, Aggression, and Attention Problems achieving salient factor pattern coefficients on an Externalizing factor and Anxiety, Depression, Withdrawal, Atypicality, and Somatization achieving salient factor pattern coefficients on an Internalizing factor. However, Learning Problems was aligned with Internalizing and had a nearly salient factor pattern coefficient. As observed in analyses with Clinical and Adaptive scales, all BASC–3 TRS Child Clinical scales, except Somatization and Anxiety, showed fair to good first unrotated factor structure coefficients using Kaufman's (1994) criteria, indicating general factor saturation.

TRS-Child Clinical Scales Second-Order EFA With SL Transformation

Due to statistically significant and moderate to large factor correlations for the four-factor model (Table 6), second-order EFA of these factor correlations was performed and results subjected to the SL orthogonalization procedure. Table 7 presents the resulting variance decomposition and metrics to assess first- and second-order factor interpretability. The second-order general factor accounted for 37.2% of total variance and 55.8% of common variance while the four first-order factors (Externalizing, School Problems, Social Disengagement, Internalizing) accounted for considerably less unique variance (2.1% to 13.4% of total variance, 3.2% to 20.0% common variance). Omega coefficients were estimated based on the decomposed variance estimates in Table 7 and the ω_H coefficient for the general factor (.766) was high and sufficient for confident scale interpretation of a unit-weighted score based on all BASC–3 TRS Child Clinical scales, but the ω_{HS} coefficients for the Externalizing, School Problems, Social Disengagement, and Internalizing factors (.118-.468) did not reach the minimum criterion for confident interpretation of unit-weighted scores based on specified indicators. *H* indices indicated an optimally weighted composite score for a general factor would account for 87.1% of variance, but the three first-order factors were not well defined by their optimally weighted indicators (*H*s < .70). The PUC estimate (.822) suggested essential unidimensionality while the ECV estimate (.558) did not. Another assessment of possible interpretation was indicated by the scale specificity estimates in Table 7 that illustrated Learning Problems, Withdrawal, Anxiety, and Somatization scales contained ample unique variance, while the Hyperactivity, Attention Problems, and Atypicality scales contained adequate unique variance for possible separate interpretation. Aggression, Conduct Problems, Attention Problems, and Depression scales did not have adequate specificity for separate interpretation.

BASC–3 TRS Adolescent

Clinical and Adaptive Scales EFA

Bartlett's Test of Sphericity, $\chi^2(105) = 9{,}127.71$, $p < .0001$; indicated the correlation matrix was not random and the KMO Measure of Sampling Adequacy of .906 far exceeded the minimum standard for factorability. Initial communality estimates ranged from .428 to .880 (*Mdn* = .780). Factor extraction criteria (see Table 1) indicated that the BASC–3 TRS Adolescent Clinical and Adaptive scales might produce two to five factors. Thus, extraction began with five factors and iteratively reduced extractions by one to examine structures.

Clinical and Adaptive scales EFA with 5 through 2 extractions produced inadequate results producing single scale (Atypicality) factor association (5 factors), salient scale factor pattern coefficients on multiple factors (cross-loading; 5, 4 [Table A8], and 3 factors), and one scale (Withdrawal) with no salient factor pattern coefficients on any factor (2 factors). Like the BASC–3 TRS Preschool and BASC–3 TRS Child, structural inadequacy observed in BASC–3 TRS Adolescent EFA including both Clinical and Adaptive scales may be the result of including both positive and negative scales in the same analyses. A final observation was that all BASC–3 TRS Adolescent Clinical and Adaptive scales, except Somatization, showed fair to good first unrotated factor structure coefficients using Kaufman's (1994) criteria, indicating saturation of a general factor.

Clinical Scales EFA

Bartlett's Test of Sphericity, $\chi^2(45) = 4{,}605.89$, $p < .0001$; indicated the correlation matrix was not random and the KMO Measure of Sampling Adequacy of .848 far exceeded the minimum standard for factorability. Initial communality estimates ranged from .420 to .810 (*Mdn* = .701). Factor extraction criteria (see Table 1) indicated that the BASC–3 TRS Adolescent Clinical scales might produce two to four factors. Thus, extraction began with four factors and iteratively reduced extractions by one to examine structures.

The most plausible factor structure was a four factor extraction (see Table 8) where simple structure was attained and Aggression, Conduct Problems, and Hyperactivity had salient factor pattern coefficients on the Externalizing factor; Anxiety, Somatization, and Depression had salient factor pattern coefficients on the Internalizing factor; Attention Problems and Learning Problems had salient factor pattern coefficients on the School Problems factor, and Withdrawal and Atypicality had salient factor pattern coefficients on the Social Disengagement factor. Extraction of three factors (see Table 9) also produced a plausible model attaining simple structure similar to the four-factor model but where Somatization and Atypicality joined the Internalizing factor. However, factor structure coefficients for Withdrawal and Atypicality on the Social Disengagement factor (see Table 8) were higher than those on the Internalizing factor (see Table 9). Extracting two factors (Table A9) produced simple structure but a less plausible structure similar to the three-factor model but with Attention Problems and Learning Problems joining the Externalizing factor. Finally, all BASC–3 TRS Adolescent Clinical scales showed fair to good first unrotated factor structure coefficients using Kaufman's (1994) criteria, indicating saturation of a general factor.

TRS-Adolescent Clinical Scales Second-Order EFA With SL Transformation

Four Group Factors. Due to statistically significant and moderate to large factor correlations for the four-factor model (Table 8), second-order EFA of these factor correlations was performed and the SL orthogonalization procedure applied to results. Table 10 presents the resulting variance decomposition and metrics to assess first- and second-order factor interpretability. The second-order general factor accounted for 45.9% of total variance and 63.2% of common variance while the four first-order factors (Externalizing, Internalizing, School Problems, Social Disengagement) accounted for considerably less unique variance (3.9% to 11.7% of total variance, 5.4% to 16.1% common variance). Omega coefficients were estimated based on the decomposed variance estimates in Table 11 and the ω_H coefficient for the general factor (.829) was high and sufficient for confident scale interpretation of a unit-weighted score based on all BASC–3 TRS Adolescent Clinical scales, but the ω_{HS} coefficients for the Externalizing, Internalizing, School Problems, Social Disengagement factors (.198-.416) did not reach the minimum criteria for confident interpretation of unit-weighted scores based on specified indicators. *H* indices indicated an optimally weighted composite score for a general factor would account for 90.3% of variance, but the three first-order factors were not well defined by their optimally weighted indicators (*H*s < .70). The general factor ECV estimate (.632) was not sufficiently high to suggest unidimensionality, but the PUC estimate (.822) did suggest the general factor to be primarily unidimensional using tentative criteria (Reise, Scheines et al., 2013). Another assessment of potential interpretation indicated by the scale specificity estimates in Table 11 illustrated the Somatization scale contained ample unique variance while the Conduct Problems, Hyperactivity, Anxiety, Attention Problems, Learning Problems, Withdrawal, and Atypicality scales contained adequate unique variance for possible separate interpretation. Aggression, Conduct Problems, Depression, and Attention Problems scales did not have adequate specificity for separate interpretation.

Three Group Factors. Due to statistically significant and moderate to large factor correlations for the three-factor model (Table 9), second-order EFA of these factor correlations was performed and results subjected to the SL orthogonalization procedure. Table 11 presents the resulting variance decomposition and metrics to assess first- and second-order factor interpretability. The second-order general factor accounted for 43.8% of total variance and 62.6% of common variance while the three first-order factors (Internalizing, Externalizing, School Problems) accounted for considerably less unique variance (4.1% to 13.1% of total variance, 5.9% to 18.7% common variance). Omega coefficients were estimated based on the decomposed variance estimates in Table 11 and the ω_H coefficient for the general factor (.770) was high and sufficient for confident scale interpretation of a unit-weighted score based on all BASC–3 TRS Adolescent Clinical scales, but the ω_{HS} coefficients for the Internalizing, Externalizing, and School Problems factors (.217-.388) did not reach the minimum threshold for confident interpretation of unit-weighted scores based on specified indicators. *H* indices indicated an optimally weighted composite score for a general factor would account for 89.9% of variance, but the three first-order factors were not well defined by their optimally weighted indicators (*H*s < .70). ECV and PUC estimates for the general factor did not suggest essential unidimensionality; however, the combination of the general factor PUC < .8 but ECV > .6 and ω_H > .7 might suggest the general factor to be primarily unidimensional (Reise, Scheines, et al., 2013). Assessment of potential interpretation based on the scale specificity estimates in Table 11 illustrated the Anxiety, Withdrawal, and Somatization scales contained ample unique variance while the Atypicality, Conduct Problems, Hyperactivity, Attention Problems, and Learning Problems scales contained adequate unique variance for possible separate interpretation.

Aggression, Conduct Problems, Depression, and Attention Problems scales did not have adequate specificity for separate interpretation. Reducing the number of extracted factors to three resulted in increasing specificity estimates for most scales.

Discussion

Throughout the development and use of the BASC, there has been a need for independent structural validity evidence to support the interpretation and use of this assessment tool. Given the relatively limited information in the BASC–3 *Manual* regarding methods and results for BASC–3 TRS construct validity based on internal structure (Messick, 1995), the present study independently examined the BASC–3 TRS Preschool, Child, and Adolescent scales using best practices in EFA (Watkins, 2018) with available summary statistics. This study is the first known independent evaluation of the factor structure of *any* of the BASC versions. While EFA would ideally begin at the item level for measures of psychopathology, this was not possible given the lack of accessible data. Thus, scale correlation matrices provided in the BASC–3 *Manual* served as input for EFA.

In EFA of BASC–3 TRS Preschool, Child, and Adolescent versions using Clinical *and* Adaptive scales, numerous problems were observed with some Clinical and Adaptive scales having salient (but inverse) cross-loadings on the same factor. This was observed with the Attention Problems scale in the BASC–3 TRS Preschool (see Tables A1 and A2); Leadership, Withdrawal, Attention Problems, and Functional Communication scales in the BASC–3 TRS Child (see Tables A4 and A5); and the Withdrawal, Leadership, and Functional Communication scales in the BASC–3 TRS Adolescent (see Table A6). This might suggest a bipolar nature of Clinical scales (problem behaviors) and Adaptive scales (positive or functional behaviors) in that they may to some extent be measuring opposite ends of a broader continuum.

In EFA of BASC–3 TRS Preschool, Child, and Adolescent versions using only Clinical scales, results provided general support for the structure of the Clinical scales. Interestingly, EFA consistently identified an *additional* factor beyond those presented in the BASC–3 *Manual*. As previously noted, the BASC–3 TRS Withdrawal and Atypicality scales for the Preschool, Child, and Adolescent versions are not used in scoring factor based composite scores (i.e., Externalizing, Internalizing, School Problems) but are included in the BSI. However, the most plausible EFA models for the BASC–3 TRS Preschool, Child, and Adolescent Clinical scales (see Tables 2, 6, and 8) identified a new factor tentatively named Social Disengagement that contained the Withdrawal and Atypicality scales. When extracting one less factor, Withdrawal and Atypicality then saliently loaded the Internalizing factor in the Preschool and Adolescent versions, but in the Child version, Atypicality saliently loaded the School Problems factor while Withdrawal had split alignment with both School Problems and Internalizing factors. There was no discussion in the BASC–3 *Manual* about the possible emergence of another factor represented by Withdrawal and Atypicality in the TRS, but this could be the result of CFA and EFA methods employed with the combined Clinical *and* Adaptive scales and with the combined normative *and* clinical samples.

While BASC–3 TRS Preschool, Child, and Adolescent Clinical scales were properly associated with their theoretical factors and Withdrawal and Atypicality appear to represent a newly identified dimension, the statistically significant correlations among the factors necessitated second-order EFA and application of the Schmid and Leiman (1957) procedure to disentangle sources of variance in order to determine how much variance was uniquely reflected by the second-order and first-order factors and to assess model-based validity. To further illustrate sources of variance among the BASC-3 TRS Clinical scales, Figures A1, A2, and A3 graphically depict proportions of general, factor, specificity, and error variance (see supplemental materials). Results indicated that none of the first-order factors (Externalizing, Internalizing, Social Disengagement, or School Problems) in the BASC–3 TRS scales contained sufficient unique true score variance for confident interpretation based on ω_{HS} coefficients as all were less than the .50 minimum criterion (Reise, 2012; Reise, Bonifay, et al., 2013) and *H* coefficients which were below the .80 criterion (Rodriguez et al., 2016a, b). The general second-order "problem behavior" factor accounted for more unique variance than all first-order factors combined in the BASC-3 TRS Preschool, Child, and Adolescent versions. These model-based indices could have been made available in the BASC–3 *Manual* based on second-order EFA from the results of oblique rotations or CFA procedures that examined higher-order (or possibly bifactor) structure, not just the oblique models. Without these types of analyses in the BASC, BASC-2, or BASC–3 for comparison, it is difficult to place the present results in a broader context.

While there may appear to be similarity between the present study second-order general "problem behavior" factor and the BASC-3 TRS BSI, they are not identical. The second-order general factor in the present study is derived from covariance among the extracted first-order factors and a hierarchically ordered construct, but the BSI is a composite score composed of the BASC-3 TRS Clinical scales "that best measure a general problem factor" (Reynolds & Kamphaus, 2015b), however, the method and criterion for selection was not reported. The BSI might be an estimate of such a general factor but there was no theoretical rationale presented for such a construct. Thus, interpretation of such a score may be questionable.

In contrast to present results with the BASC-3, a rival teacher report behavior-rating scale of youth psychopathology developed around the same time as the original BASC demonstrated more favorable results. The Adjustment Scales for Children and Adolescents (ASCA; McDermott et al., 1993) was co-normed with the Differential Abilities Scale (DAS; Elliott,

1990) by The Psychological Corporation and correlations with the ASCA standardization sample revealed much lower syndrome (scale) relationships (highest were .49 between Attention-Deficit/Hyperactive and Solitary Aggressive-Provocative and .49 between Oppositional Defiant and Solitary Aggressive-Provocative). Further, the obliquely rotated Overactivity and Underactivity syndromes (similar to Externalizing and Internalizing dimensions) have been repeatedly observed to be essentially independent in the standardization sample (McDermott, 1993) and in various independent samples with obliquely rotated factor correlations < .22 (r = .08 [Canivez, 2004], r = -.02 [Canivez, 2006], r = .06 [Canivez & Bohan, 2006], r = 0 [Canivez & Beran, 2009], r = .21 [Canivez & Sprouls, 2010]. Thus, a second-order (higher-order) EFA to explicate some "general problem" factor was unnecessary as there was no appreciable factor covariance among the Overactivity and Underactivity syndromes (Tabachnick & Fidell, 2007) and true score variance was a result of ASCA Overactivity and its core syndromes specificity and the separate ASCA Underactivity and its core syndromes specificity. In the BASC-3 TRS, observed true score variance was related to *three* sources (general factor, first-order factors, and scales). Even if this general dimension remains undefined, ASCA results illustrate that it is not inevitable that behavior-rating scales contain substantial covariance among scales and global externalizing and internalizing factors that would require explication of a higher-order general dimension. It may very well be that such a factor merely represents a diminutive "*g*" (Stankov, 2002) that is locally defined.

Benson et al. (2018) noted "a test-derived score with interpretive relevance (a) provides a good representation of the construct targeted for measurement, (b) is distinct from conceptually similar constructs, (c) is likely to be replicable across data sets and methods, and (d) has adequate unique, reliable variance such that it is statistically distinguishable from test-derived scores reflecting conceptually similar constructs" (p. 1030). Results of the present study seem to indicate that while a, b, and c might be satisfied, there appear to be problems regarding d where it was observed that first-order (group) factors appear to contain too little unique portions of variance for confident interpretation.

Scale specificity estimates, unique reliable variance within individual BASC–3 TRS scales, indicated that many BASC–3 TRS scales (but not all) contained ample or adequate variance for possible interpretation according to Kaufman and Lichtenberger (2005) criteria. Whether these portions of unique variance are useful must be assessed using methods comparing them against external criteria. Use of incremental validity (Hunsley, 2003) and diagnostic utility (Kessel & Zimmerman, 1993) methods will be particularly valuable in such assessments. Cluster analyses (i.e., DiStefano et al., 2010; McDermott & Weiss, 1995) or latent profile analyses (LPA; i.e., Flaherty & Kiff, 2012) might also help provide assessment of BASC–3 TRS scale utility. A recent diagnostic utility study by Zhou et al. (2020), however, found the BASC–3 Autism Probability Index from the Parent Rating Scales (PRS) and Teacher Rating Scales (TRS) to differentiate individuals with Autism from those with ADHD with Receiver Operating Characteristic Curve (AUC) estimates of .85 and .83, respectively. Combinations of the Developmental Social Disorders content scale and the Atypicality and Withdrawal clinical scales for the PRS and TRS also produced AUC estimates of .86 and .84, respectively, in differentiating individuals with Autism from those with ADHD.

Limitations and Future Directions

The primary limitation of the present study was the inability to begin EFA at the item level using item polychoric correlations and examining univariate and multivariate normality estimates. Polychoric correlations may be more appropriate in the case of ordinal (graded-response) items (Watkins, 2018) and produce higher (or lower) item correlations that might produce more optimal first-order, second-order, and possibly third-order structural validity results and better estimates of item covariance. Better estimates might also have resulted in different estimates of unique contributions of scales and second-order factors better supporting interpretation. In the present study it was only possible to test the structural validity of the BASC–3 TRS using *scale* intercorrelations, which would be equivalent to examining the structural validity of an intelligence test by starting factor analysis with the factor scores rather than the subtests. Unfortunately, without access to the BASC–3 TRS *item* raw score data this is the best that can be done using standardization summary statistics available in the BASC–3 *Manual*.

Another limitation is that precise estimates of scale skewness and kurtosis are unknown so their effects on the scale intercorrelations that served as the basis of EFA is also unknown. While skewness and kurtosis might impact the correlations, the method of factor extraction (principal axis) used does not have distributional assumptions and is commonly used with nonnormally distributed data. A final limitation is that the present study examined EFA of the normative sample correlations and not the full sample of normative *and* clinical samples as were reported in the BASC–3 *Manual*. This was not possible because the BASC–3 *Manual* separately presents correlations for the normative sample and the clinical sample and a correlation matrix of the full combined sample was not available. Thus, results may differ from those reported in the BASC–3 *Manual.*

Item and scale level raw data are necessary to perform EFA and CFA according to best practices. Thus, it is incumbent on the publisher to conduct such analyses and fully disclose results using best EFA practices to elucidate these important metrics. While Appelbaum et al. (2018) published important reporting standards necessary for journal publications, these same standards should also apply to test publishers in full disclosure of critical psychometric properties in test manuals. Full accounting and disclosure of psychometric details following best practices are also encouraged in the *Standards for Educational and Psychological Testing* (AERA et al., 2014) to provide the necessary evidence to support all interpretations of test scores and comparisons.

While the present study examined, supported, and extended some BASC–3 TRS construct validity using EFA, best practices in CFA should follow-up these results to test various models in a confirmatory framework. Such analyses *should* begin with items as measured variables or indicators so that proper CFA may be conducted. Without *item* raw data, any independent CFA of the BASC–3 TRS standardization sample would need to start with the *scale* correlation matrix which is less than ideal.

Conclusion

Present results indicated the BASC–3 TRS Preschool, Child, and Adolescent Clinical scales construct validity based on internal structure was largely supported by proper alignment with their theoretical composite score dimension. The discovery of a possible new factor, Social Disengagement, which was consistently loaded by Withdrawal and Atypicality in all three TRS forms, suggests that an additional composite score that combines these two BASC–3 scales might be considered. However, it was also found that when considering second-order EFA to structurally explicate first-order factor correlations and partitioning variance to first- and second-order factors using the Schmid and Leiman (1957) procedure, the BASC–3 TRS Preschool, Child, and Adolescent factors represented by BASC-3 composite scores (Externalizing, Internalizing, School Problems, and Social Disengagement) did not contain sufficient portions of unique variance for confident interpretation. Clinical scale specificity (unique true score variance within the scale) was ample or adequate for many (but not all) of the BASC–3 TRS scales, which if identified as useful through other methods (incremental validity, diagnostic utility, latent profile analysis), might offer additional interpretive value. While the present study examined construct validity of the BASC-3 TRS Clinical and Adaptive scales, it did not examine construct validity of content scales or other indexes, so BASC-3 users will need to consult the extent literature to determine viability of such scores in clinical use. Clinicians are encouraged to consider the information provided by the present investigation to follow Weiner's (1989) advice for ethical test score interpretation to "(a) know what their tests can do and (b) act accordingly" (p. 829). It seems necessary that the publisher conduct a reanalysis of structural validity using best practices in EFA and CFA methods suggested in the present study beginning with items as measured variables and appropriate item level correlational estimates. Full explication of all structural analyses and related assessment of viability of all scores is needed to better understand the veracity of BASC-3 TRS scores.

References

American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (2014). *Standards for educational and psychological testing* (3rd ed.). American Educational Research Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Americans With Disabilities Act of 1990, 42 U.S.C.A. §§ 12101-12213 (1990).

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3-25. http://dx.doi.org/10.1037/amp0000191

Bartlett, M. S. (1954). A further note on the multiplying factors for various chi-square approximations in factor analysis. *Journal of the Royal Statistical Society: Series B*, *16*(2), 296-298. https://doi.org/10.1111/j.2517-6161.1954.tb00174.x

Beaujean, A. A. (2015). Adopting a new test edition: Psychometric and practical considerations. *Research and Practice in the Schools*, *3*(1), 51-57.

Beaujean, A. A., & Benson, N. F. (2019). Theoretically-consistent cognitive ability test development and score interpretation. *Contemporary School Psychology, 23,* 126–137. https://doi.org/10.1007/s40688-018-0182-1

Benson, N. F., Beaujean, A. A., McGill, R. J., & Dombrowski, S. C. (2018). Revisiting Carroll's survey of factor-analytic studies: Implications for the clinical assessment of intelligence. *Psychological Assessment, 30*(8), 1028-1038. http://dx.doi.org/10.1037/pas0000556

Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefer, S. A., & Morgan, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 national survey. *Journal of School Psychology, 72*, 29-48. https://doi.org/10.1016/j.jsp.2018.12.004

Brown, T. A. (2014). *Confirmatory factor analysis for applied research.* Guilford Press.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*(4), 796-846. https://doi.org/10.1111/j.1467-6494.2011.00749.x

Canivez, G. L. (2004). Replication of the Adjustment Scales for Children and Adolescents core syndrome factor structure. *Psychology in the Schools, 41*(2), 191-199. https://doi.org/10.1002/pits.10121

Canivez, G. L. (2006). Adjustment Scales for Children and Adolescents and Native American Indians: Factorial validity generalization for Ojibwe youths. *Psychology in the Schools, 43*(6), 685-694. https://doi.org/10.1002/pits.20179

Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247-271). Hogrefe.

Canivez, G. L., & Beran, T. N. (2009). Adjustment Scales for Children and Adolescents: Factorial validity in a Canadian sample. *Canadian Journal of School Psychology, 24*(4), 284-302. https://doi.org/10.1177/0829573509344344

Canivez, G. L., & Bohan, K. (2006). Adjustment Scales for Children and Adolescents and Native American Indians: Factorial validity generalization for Yavapai Apache youths. *Journal of Psychoeducational Assessment, 24*(4), 329-341. https://doi.org/10.1177/0734282906291397

Canivez, G. L., & Sprouls, K. (2010). Adjustment Scales for Children and Adolescents: Factorial validity generalization with Hispanic/Latino youths. *Journal of Psychoeducational Assessment, 28*(3), 209-221. https://doi.org/10.1177/0734282909349213

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219-251. http://doi.org/10.1111/j.14676494.2011.00739.x

Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, *70*(6), 885-901. http://doi.org/10.1177/0013164410379332

DiStefano, C. A., Kamphaus, R. W., & Mindrila, D. L. (2010). Typology of teacher-rated child behavior: Revisiting subgroups over 10 years later. *School Psychology Quarterly, 25*(3), 152-163. https://doi.org/10.1037/a0020913

Elliott, C. D. (1990). *Differential Ability Scales: Introductory and technical handbook*. Psychological Corporation.

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.

Flaherty, B. P., & Kiff, C. J. (2012). Latent class and latent profile models. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *Handbook of research methods in psychology: Volume 1 Foundations, planning, measures, and psychometrics* (pp. 391–404). American Psychological Association. https://doi.org/10.1037/13621-019

Frazier, T. W., & Youngstrom, E. A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence*, *35*(2), 169-182. http://doi.org/10.1016/j.intell.2006.07.002

Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, *48*(5), 639-662. http://doi.org/10.1080/00273171.2013.804398

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*(3), 377-393. http://doi.org/10.1177/0013164495055003002

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Lawrence Erlbaum.

Gorsuch, R. L. (2003). Factor analysis. In J. A. Schinka & F. F. Velicer (Eds.), *Handbook of psychology: Vol. 2. Research methods in psychology* (pp. 143–164). Wiley.

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika, 19,* 149-161. https://doi.org/10.1007/BF02289162

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 195-216). Scientific Software International.

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

Hunsley, J. (2003). Introduction to the special section on incremental validity and utility in clinical assessment. *Psychological Assessment, 15*(4), 443-445. https://doi.org/10.1037/1040-3590.15.4.443

Individuals with Disabilities Education Act, 20 U.S.C.A. § 1400 et seq. (2004).

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. http://doi.org/10.1007/BF02291575

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73. https://doi.org/10.1111/jedm.12000

Kaufman, A. S. (1994). *Intelligent testing with the WISC–III*. Wiley.

Kaufman, A. S., & Lichtenberger, E. O. (2005). *Assessing adolescent and adult intelligence* (3rd Ed.). Wiley.

Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment, 5*(4), 395–399. https://doi.org/10.1037/1040-3590.5.4.395.

Kline, P. (1994). *An easy guide to factor analysis.* Routledge.

Li, C-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust likelihood and weighted least squares. *Behavioral Research methods, 48*, 936-949. https://doi.org/10.3758/s13428-015-0619-7

McDermott, P. A. (1993). National standardization of uniform multisituational measures of child and adolescent behavior pathology. *Psychological Assessment, 5*(4), 413-424. https://doi.org/10.1037/1040-3590.5.4.413

McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment Scales for Children and Adolescents.* Edumetric and Clinical Science. https://doi.org/10.1037/t00682-000

McDermott, P. A., & Weiss, R. V. (1995). A normative typology of healthy, subclinical, and clinical behavior styles among American children and adolescents. *Psychological Assessment, 7*(2), 162-170. https://doi.org/10.1037/1040-3590.7.2.162

McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9)*,* 741-749. https://doi.org/10.1037/0003-066X.50.9.741

Nasser, F., Benson, J., & Wisenbaker, J. (2002). The performance of regression-based variations of the visual scree for determining the number of common factors. *Educational and Psychological Measurement*, *62*(3), 397-419. http://doi.org/10.1177/00164402062003001

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396-402. http://doi.org/10.3758/BF03200807

Rehabilitation Act, 29 U.S.C.A. §. 701 et seq. (1973).

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95*(2)*,* 129-140. https://doi.org/10.1080/00223891.2012.725437

Reise, S. P., Scheines, R., Widaman, K. F., and Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement, 73*(1), 5-26. https://doi.org/10.1177/0013164412449831

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children*. American Guidance Service.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children, second edition*. Pearson

Reynolds, C. R., & Kamphaus, R. W. (2015a). *Behavior Assessment System for Children, third edition*. Pearson.

Reynolds, C. R., & Kamphaus, R. W. (2015b). *Behavior Assessment System for Children, third edition: Manual*. Pearson.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2)*,* 137-150. https://doi.org/10.1037/met0000045

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3)*,* 223-237. https://doi.org/10.1080/00223891.2015.1089249

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*(1), 53-61. http://doi.org/10.1007/BF02289209

Schreiber, J. B., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, *99*(6), 323-337. https://doi.org/10.3200/JOER.99.6.323-338

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31*(12), 1428-1441. https://doi.org/10.1037/pas0000623

Stankov, L. (2002). *g*: A diminutive general. In R. J. Sternberg & E. L. Grigorenko (Eds.), The general factor of intelligence: How general is it? (pp. 19-37). Lawrence Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* American Psychological Association.

Thurstone, L. L. (1947). *Multiple-factor analysis*. University of Chicago Press.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321-327. http://doi.org/10.1007/BF02293557

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin, & E. Helms (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41-71). Springer. https://doi.org/10.1007/978-1-4615-4397-8_3

Watkins, M. W. (2000). *Monte Carlo PCA for Parallel Analysis* [Computer Software]. Ed & Psych Associates.

Watkins, M. W. (2007). *SEscree* [Computer software]. Ed & Psych Associates.

Watkins, M. W. (2013). *Omega* [Computer software]. Ed & Psych Associates.

Watkins, M. W. (2017). The reliability of multidimensional neuropsychological measures: From alpha to omega. *The Clinical Neuropsychologist, 31*(6-7), 1113-1126. https://doi.org/10.1080/13854046.2017.1317364

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology, 44*(3), 219-246. https://doi.org/10.1177/0095798418771807

Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment, 53*(4), 827-831. https://doi.org/10.1207/s15327752jpa5304_18

Wolff, H.-G., & Preising, K. (2005). Exploring item and higher order factor structure with the Schmid-Leiman solution: Syntax codes for SPSS and SAS. *Behavior Research Methods, 37*(1), 48-58. https://doi.org/10.3758/BF03206397

Zhou, X., Reynolds, C., Zhu, J., & Kamphaus, R. W. (2020). Differentiating autism from ADHD in children and adolescents. *Journal of Pediatric Neuropsychology, 6*, 61-65. https://doi.org/10.1007/s40817-020-00082-7

Zoski, K. W., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement*, *56*(3), 443-451. https://doi.org/10.1177/0013164496056003006

Table 1

Number of BASC–3 Teacher Rating Scale Factors Suggested for Extraction Across Five Different Extraction Indicators for Preschool, Child, and Adolescent Versions

	BASC–3 Clinical and Adaptive Scales		
Extraction Indicator	Preschool	Child	Adolescent
Eigenvalue > 1	3	3	3
Scree Test (Visually Examined)	4-5	2-5	2-5
Standard Error of Scree (SE_{Scree})	3	5	4
Horn's Parallel Analysis (HPA)	3	3	3
Minimum Average Partials (MAP)	3	4	4
Publisher (Theory) Proposed	3	4	4
	BASC–3 Clinical Scales		
Extraction Indicator	Preschool	Child	Adolescent
Eigenvalue > 1	2	3	2
Scree Test (Visually Examined)	2-4	2-4	2-4
Standard Error of Scree (SE_{Scree})	2	3	2
Horn's Parallel Analysis (HPA)	2	2	2
Minimum Average Partials (MAP)	2	2	2
Publisher (Theory) Proposed	2	3	3

Note. BASC–3 = Behavior Assessment System for Children-Third Edition

Table 2
BASC–3 Teacher Rating Scale-Preschool Exploratory Factor Analysis of Clinical Scales: Three Oblique Factor Solution for the Standardization General Norm Sample (N = 500)

	General	F1: Externalizing		F2: Internalizing		F3: Social Disengagement		
BASC–3 Scale	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Hyperactivity	.781	**1.039**	.960	-.128	.371	-.030	.437	.936
Aggression	.795	**.762**	.838	.271	.587	-.113	.433	.748
Attention Problems	.639	**.632**	.697	-.231	.277	*.349*	.547	.567
Anxiety	.626	-.151	.323	**.782**	.799	.164	.523	.662
Depression	.837	*.319*	.664	**.703**	.857	-.009	.550	.810
Somatization	.378	-.088	.194	**.548**	.515	.018	.279	.270
Withdrawal	.640	-.083	.382	.184	.545	**.721**	.780	.631
Atypicality	.730	.189	.560	.067	.525	**.652**	.787	.654
Eigenvalue		7.43		2.72		.92		
% Variance		51.30		17.60		4.55		
Promax Based Factor Correlations		F1		F2		F3		
F1: Externalizing		–						
F2: Internalizing		.497		–				
F3: Social Disengagement		.518		.559		–		

Note. S = Structure Coefficient, *P* = Pattern Coefficient, h^2 = Communality (Extraction). General structure coefficients are based on the first unrotated factor coefficients (*g* loadings). Salient pattern coefficients presented in bold (pattern coefficient ≥ .40) and aligned (.30-.39) in italic.

Table 3

BASC–3 Teacher Rating Scale-Preschool Exploratory Factor Analysis of Clinical Scales: Two Oblique Factor Solution for the Standardization General Norm Sample (N = 500)

	General	F1: Externalizing		F2: Internalizing		
BASC–3 Scale	*S*	*P*	*S*	*P*	*S*	h^2
Hyperactivity	.793	**1.078**	.958	-.210	.405	.947
Attention Problems	.636	**.727**	.712	-.026	.389	.508
Aggression	.787	**.713**	.807	.164	.571	.669
Anxiety	.634	-.169	.349	**.909**	.812	.678
Depression	.827	.292	.663	**.649**	.816	.723
Withdrawal	.611	.117	.450	**.583**	.650	.432
Somatization	.380	-.116	.204	**.560**	.494	.253
Atypicality	.707	*.350*	.607	**.451**	.651	.947
Eigenvalue		4.12		1.32		
% Variance		46.95		12.01		
Promax Based Factor Correlation		F1		F2		
F1: Externalizing		–				
F2: Internalizing		.571		–		

Note. S = Structure Coefficient, *P* = Pattern Coefficient, h^2 = Communality (Extraction). General structure coefficients are based on the first unrotated factor coefficients (*g* loadings). Salient pattern coefficients presented in bold (pattern coefficient ≥ .40) and aligned (.30-.39) in italic.

Table 4
Sources of Variance in the BASC–3 Teacher Rating Scale-Preschool Clinical Scales Standardization General Norm Sample (N = 500) According to the Schmid-Leiman Transformed Higher-Order Factor Model with Three Group Factors

	General		F1: Externalizing		F2: Internalizing		F3: Social Disengagement				
BASC–3 Scale	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2	s^2
Hyperactivity	.589	.347	**.763**	**.582**	-.087	.008	-.019	.000	.937	.063	.000[1]
Aggression	.630	.397	**.559**	**.312**	.184	.034	-.073	.005	.749	.251	.151*
Attention Problems	.526	.277	**.464**	**.215**	-.157	.025	.226	.051	.568	.432	.332**
Anxiety	.596	.355	-.111	.012	**.532**	**.283**	.106	.011	.662	.338	.128
Depression	.725	.526	.234	.055	**.478**	**.228**	-.006	.000	.809	.191	.031
Somatization	.356	.127	-.065	.004	**.373**	**.139**	.012	.000	.270	.730	.550**
Withdrawal	.628	.394	-.061	.004	.125	.016	**.467**	**.218**	.632	.368	.168
Atypicality	.674	.454	.139	.019	.046	.002	**.422**	**.178**	.654	.346	.236*
Total Variance		.360		.151		.092		.058			
Explained Common Variance		.545		.228		.139		.088			
ω		.905		.878		.779		.767			
ω_H / ω_{HS}		.716		.449		.315		.244			
Factor Correlation		.846		.670		.561		.494			
H		.829		.680		.460		.331			
PUC		.750									
FDI		.911		.824		.678		.576			

Note. *b* = loading of scale on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, s^2 = scale specificity (uniqueness-error), ω_H = Omega-hierarchical (general factor), ω_{HS} = Omega-hierarchical subscale (group factors), *H* = construct replicability coefficient, PUC = percentage of uncontaminated correlations, FDI = factor determinacy index. Bold type indicates coefficients and variance estimates consistent with the theoretically proposed factor. Italic type indicates coefficients and variance estimates associated with an alternate factor (where cross-loading *b* was larger than for the theoretically assigned factor). Light shading indicates minimum standard met, dark shading indicates preferred standard met. [1]$1 - \alpha > u^2$ so s^2 set to 0. *Adequate, **Ample (Kaufman & Lichtenberger, 2005).

Table 5

Sources of Variance in the BASC–3 Teacher Rating Scale-Preschool Clinical Scales Standardization General Norm Sample (N = 500) According to the Schmid-Leiman Transformed Higher-Order Factor Model with Two Group Factors

	General		F1: Externalizing		F2: Internalizing				
BASC–3 Scale	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2	s^2
Hyperactivity	.655	.429	**.707**	**.500**	-.138	.019	.948	.052	.000[1]
Attention Problems	.529	.280	**.477**	**.228**	-.017	.000	.508	.492	.392**
Aggression	.662	.438	**.468**	**.219**	.108	.012	.669	.331	.231*
Anxiety	.559	.312	-.111	.012	**.596**	**.355**	.680	.320	.110
Depression	.710	.504	.191	.036	**.426**	**.181**	.722	.278	.118
Withdrawal	.529	.280	.077	.006	**.382**	**.146**	.432	.568	.368**
Somatization	.335	.112	-.076	.006	**.367**	**.135**	.253	.747	.567**
Atypicality	.605	.366	.230	.053	**.296**	**.088**	.507	.493	.383**
Total Variance		.340		.132		.117			
Explained Common Variance		.577		.225		.198			
ω		.891		.871		.824			
ω_H / ω_{HS}		.668		.387		.299			
Factor Correlation		.818		.622		.547			
H		.818		.612		.544			
PUC		.536							
FDI		.904		.782		.544			

Note. *b* = scale loading on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, s^2 = scale specificity (uniqueness-error), ω = Omega, ω_H = Omega-hierarchical (general factor), ω_{HS} = Omega-hierarchical subscale (group factors), *H* = construct replicability coefficient, PUC = percentage of uncontaminated correlations, FDI = factor determinacy index. Bold type indicates coefficients and variance estimates consistent with the theoretically proposed factor. Light shading indicates minimum standard met, dark shading indicates preferred standard met. [1] $1 - \alpha > u^2$ so s^2 set to 0. *Adequate, **Ample (Kaufman & Lichtenberger, 2005).

Table 6
BASC–3 Teacher Rating Scale-Child Exploratory Factor Analysis of Clinical Scales: Four Oblique Factor Solution for the Standardization General Norm Sample (N = 600)

	General	F1: Externalizing		F2: School Problems		F3: Social Disengagement		F4: Internalizing		
BASC–3 Scale	*S*	*P*	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Aggression	.795	**.953**	.902	-.191	.409	.088	.544	.014	.420	.834
Conduct Problems	.825	**.888**	.904	.044	.535	-.054	.529	.051	.414	.819
Hyperactivity	.761	**.827**	.845	.222	.593	-.113	.441	-.098	.260	.755
Attention Problems	.794	.273	.680	**.772**	.901	-.009	.571	-.061	.312	.856
Learning Problems	.586	-.128	.382	**.723**	.738	.077	.515	.124	.365	.569
Withdrawal	.580	-.095	.380	.005	.438	**.866**	.755	-.086	.453	.581
Atypicality	.743	.121	.577	.158	.606	**.682**	.792	-.083	.479	.667
Anxiety	.454	-.194	.234	.171	.324	.030	.502	**.724**	.718	.545
Somatization	.296	.074	.217	-.031	.112	-.182	.256	**.618**	.518	.286
Depression	.780	.350	.662	-.073	.436	.238	.731	**.495**	.781	.760
Eigenvalue		4.95		1.39		1.08		.77		
% Variance		46.56		9.93		6.89		3.36		
Promax Based Factor Correlations		F1		F2		F3		F4		
F1: Externalizing		–								
F2: School Problems		.570		–						
F3: Social Disengagement		.589		.597		–				
F4: Internalizing		.432		.339		.668		–		

Note. S = Structure Coefficient, *P* = Pattern Coefficient, h^2 = Communality (Extraction). General structure coefficients are based on the first unrotated factor coefficients (*g* loadings). Salient pattern coefficients presented in bold (pattern coefficient ≥ .40) and aligned (.30-.39) in italic.

Table 7

Sources of Variance in the BASC–3 Teacher Rating Scale-Child Clinical Scales Standardization General Norm Sample (N = 600) According to the Schmid-Leiman Transformed Higher-Order Factor Model with Four Group Factors

	General		F1: Externalizing		F2: School Problems		F3: Social Disengagement		F4: Internalizing				
BASC–3 Scale	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2	s^2
Aggression	.635	.403	**.675**	**.456**	-.142	.020	.035	.001	.011	.000	.880	.120	.020
Conduct Problems	.639	.408	**.629**	**.396**	.033	.001	-.022	.000	.039	.002	.807	.193	.093
Hyperactivity	.565	.319	**.586**	**.343**	.165	.027	-.045	.002	-.075	.006	.697	.303	.243*
Attention Problems	.660	.436	.193	.037	**.575**	**.331**	-.004	.000	-.047	.002	.806	.194	.134
Learning Problems	.542	.294	-.091	.008	**.539**	**.291**	.031	.001	.095	.009	.603	.397	.307**
Withdrawal	.675	.456	-.067	.004	.004	.000	**.345**	**.119**	-.066	.004	.584	.416	.276**
Atypicality	.763	.582	.086	.007	.118	.014	**.272**	**.074**	-.064	.004	.682	.318	.178*
Anxiety	.469	.220	-.137	.019	.127	.016	.012	.000	**.556**	**.309**	.564	.436	.296**
Somatization	.261	.068	.052	.003	-.023	.001	-.073	.005	**.474**	**.225**	.301	.699	.579**
Depression	.734	.539	.248	.062	-.054	.003	.095	.009	**.380**	**.144**	.757	.243	.093
Total Variance		.372		.134		.070		.021		.071			
ECV		.558		.200		.105		.032		.106			
ω		.922		.912		.805		.761		.734			
ω_H / ω_{HS}		.766		.468		.372		.118		.353			
Factor Correlation		.875		.684		.610		.344		.594			
H		.871		.668		.465		.177		.475			
PUC		.822											
FDI		.933		.817		.689		.421		.689			

Note. *b* = scale loading on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, s^2 = scale specificity (uniqueness-error), ECV = explained common variance, ω = Omega, ω_H = Omega-hierarchical (general factor), ω_{HS} = Omega-hierarchical subscale (group factors), *H* = construct replicability coefficient, PUC = percentage of uncontaminated correlations, FDI = factor determinacy index. Bold type indicates coefficients and variance estimates consistent with the theoretically proposed factor. Light shading indicates minimum standard met, dark shading indicates preferred standard met. *Adequate, **Ample (Kaufman & Lichtenberger, 2005).

Table 8

BASC–3 Teacher Rating Scale-Adolescent Exploratory Factor Analysis of Clinical Scales: Four Oblique Factor Solution for the Standardization General Norm Sample (N = 600)

	General	F1: Externalizing		F2: Internalizing		F3: School Problems		F4: Social Disengagement		
BASC–3 Scale	*S*	*P*	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Aggression	.788	**1.033**	.925	-.050	.485	-.140	.544	.026	.493	.867
Conduct Problems	.799	**.907**	.909	-.047	.478	.083	.643	-.050	.472	.832
Hyperactivity	.781	**.773**	.856	.064	.496	.188	.666	-.142	.450	.755
Anxiety	.613	-.111	.385	**.850**	.784	.041	.410	-.031	.602	.622
Somatization	.565	-.025	.385	**.770**	.694	.050	.389	-.112	.511	.486
Depression	.843	.256	.667	**.537**	.878	-.121	.526	*.329*	.832	.852
Attention Problems	.799	.223	.725	-.079	.500	**.782**	.908	.034	.532	.851
Learning Problems	.722	-.064	.560	.129	.554	**.767**	.825	.057	.549	.700
Withdrawal	.599	-.162	.360	-.050	.608	.051	.422	**.912**	.810	.672
Atypicality	.783	.192	.626	.119	.704	.093	.585	**.537**	.791	.683
Eigenvalue		5.67		1.44		.74		.67		
% Variance		54.07		11.49		4.87		2.78		
Promax Based Factor Correlations		F1		F2		F3		F4		
F1: Externalizing		–								
F2: Internalizing		.572		–						
F3: School Problems		.675		.543		–				
F4: Social Disengagement		.565		.792		.555		–		

Note. *S* = Structure Coefficient, *P* = Pattern Coefficient, h^2 = Communality (Extraction). General structure coefficients are based on the first unrotated factor coefficients (*g* loadings). Salient pattern coefficients presented in bold (pattern coefficient ≥ .40) and aligned (.30-.39) in italic. Due to Heywood case in communality estimation, extraction was limited to two iterations as per Gorsuch (2003).

Table 9

BASC–3 Teacher Rating Scale-Adolescent Exploratory Factor Analysis of Clinical Scales: Three Oblique Factor Solution for the Standardization General Norm Sample (N = 600)

	General	F1: Internalizing		F2: Externalizing		F3: School Problems		
BASC–3 Scale	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Depression	.843	**.850**	.912	.237	.645	-.124	.542	.861
Anxiety	.608	**.792**	.740	-.091	.368	.001	.413	.553
Withdrawal	.588	**.733**	.709	-.171	.337	.124	.448	.518
Somatization	.559	**.644**	.645	-.007	.371	.009	.388	.416
Atypicality	.780	**.606**	.780	.163	.603	.134	.605	.652
Aggression	.790	.019	.520	**1.003**	.925	-.133	.550	.864
Conduct Problems	.801	-.058	.506	**.887**	.911	.085	.645	.833
Hyperactivity	.783	-.038	.507	**.759**	.856	.178	.664	.749
Attention Problems	.801	-.060	.544	.219	.719	**.799**	.910	.851
Learning Problems	.722	.159	.585	-.059	.549	**.770**	.825	.696
Eigenvalue		5.67		1.44		.74		
% Variance		53.90		11.20		4.84		
Promax Based Factor Correlations		F1		F2		F3		
F1: Internalizing		–						
F2: Externalizing		.578		–				
F3: School Problems		.597		.670		–		

Note. *S* = Structure Coefficient, *P* = Pattern Coefficient, h^2 = Communality (Extraction). General structure coefficients are based on the first unrotated factor coefficients (*g* loadings). Salient pattern coefficients presented in bold (pattern coefficient ≥ .40) and aligned (.30-.39) in italic.

Table 10

Sources of Variance in the BASC–3 Teacher Rating Scale-Adolescent Clinical Scales Standardization General Norm Sample (N = 600) According to the Schmid-Leiman Transformed Higher-Order Factor Model with Four Group Factors

	General		F1: Externalizing		F2: Internalizing		F3: School Problems		F4: Social Disengagement				
BASC–3 Scale	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2	s^2
Aggression	.654	.428	**.682**	**.465**	-.028	.001	-.096	.009	.014	.000	.903	.097	.007
Conduct Problems	.661	.437	**.599**	**.359**	-.026	.001	.057	.003	-.028	.001	.800	.200	.120
Hyperactivity	.653	.426	**.510**	**.260**	.036	.001	.128	.016	-.079	.006	.710	.290	.220*
Anxiety	.626	.392	-.073	.005	**.474**	**.225**	.028	.001	-.017	.000	.623	.377	.247*
Somatization	.564	.318	-.017	.000	**.429**	**.184**	.034	.001	-.062	.004	.507	.493	.373**
Depression	.823	.677	.169	.029	**.300**	**.090**	-.083	.007	.183	.033	.836	.164	.024
Attention Problems	.701	.491	.147	.022	-.044	.002	**.534**	**.285**	.019	.000	.800	.200	.140
Learning Problems	.666	.444	-.042	.002	.072	.005	**.524**	**.275**	.032	.001	.726	.274	.154*
Withdrawal	.633	.401	-.107	.011	-.028	.001	.035	.001	**.506**	**.256**	.670	.330	.210*
Atypicality	.758	.575	.127	.016	.066	.004	.064	.004	**.298**	**.089**	.688	.312	.192*
Total Variance		.459		.117		.051		.060		.039			
ECV		.632		.161		.071		.083		.054			
ω		.947		.919		.832		.855		.792			
ω_H / ω_{HS}		.829		.416		.219		.320		.198			
Factor Correlation		.911		.450		.468		.566		.445			
H		.903		.640		.381		.437		.306			
PUC		.822											
FDI		.950		.800		.617		.661		.553			

Note. *b* = scale loading on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, s^2 = scale specificity (uniqueness-error), ECV = explained common variance, ω = Omega, ω_H = Omega-hierarchical (general factor), ω_{HS} = Omega-hierarchical subscale (group factors), *H* = construct replicability coefficient, PUC = percentage of uncontaminated correlations, FDI = factor determinacy index. Bold type indicates coefficients and variance estimates consistent with the theoretically proposed factor. Light shading indicates minimum standard met, dark shading indicates preferred standard met. *Adequate, **Ample (Kaufman & Lichtenberger, 2005).

Table 11

Sources of Variance in the BASC–3 Teacher Rating Scale-Adolescent Clinical Scales Standardization General Norm Sample (N = 600) According to the Schmid-Leiman Transformed Higher-Order Factor Model with Three Group Factors

	General		F1: Internalizing		F2: Externalizing		F3: School Problems				
BASC–3 Scale	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2	s^2
Depression	.698	.487	**.592**	**.350**	.140	.020	-.069	.005	.862	.138	.000[1]
Anxiety	.496	.246	**.551**	**.304**	-.054	.003	.001	.000	.553	.447	.317**
Withdrawal	.492	.242	**.510**	**.260**	-.101	.010	.069	.005	.517	.483	.363**
Somatization	.464	.215	**.448**	**.201**	-.004	.000	.005	.000	.416	.584	.464**
Atypicality	.678	.460	**.422**	**.178**	.096	.009	.075	.006	.653	.347	.227*
Aggression	.712	.507	.013	.000	**.594**	**.353**	-.074	.005	.865	.135	.045
Conduct Problems	.744	.554	-.040	.002	**.525**	**.276**	.047	.002	.833	.167	.087
Hyperactivity	.732	.536	-.026	.001	**.449**	**.202**	.099	.010	.748	.252	.182*
Attention Problems	.797	.635	-.042	.002	.130	.017	**.444**	**.197**	.851	.149	.089
Learning Problems	.706	.498	.111	.012	-.035	.001	**.428**	**.183**	.695	.305	.185*
Total Variance		.438		.131		.089		.041			
Explained Common Variance		.626		.187		.127		.059			
ω		.944		.875		.927		.861			
ω_H / ω_{HS}		.770		.388		.314		.217			
Factor Correlation		.877		.623		.561		.466			
H		.899		.642		.541		.320			
PUC		.689									
FDI		.948		.801		.735		.565			

Note. *b* = loading of scale on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, s^2 = scale specificity (uniqueness-error), ω_H = Omega-hierarchical (general factor), ω_{HS} = Omega-hierarchical subscale (group factors), *H* = construct replicability coefficient, PUC = percentage of uncontaminated correlations, FDI = factor determinacy index. Bold type indicates coefficients and variance estimates consistent with the theoretically proposed factor. Italic type indicates coefficients and variance estimates associated with an alternate factor (where cross-loading *b* was larger than for the theoretically assigned factor). Light shading indicates minimum standard met, dark shading indicates preferred standard met.

[1]1 – α > u^2 so s^2 set to 0. *Adequate, **Ample (Kaufman & Lichtenberger, 2005).

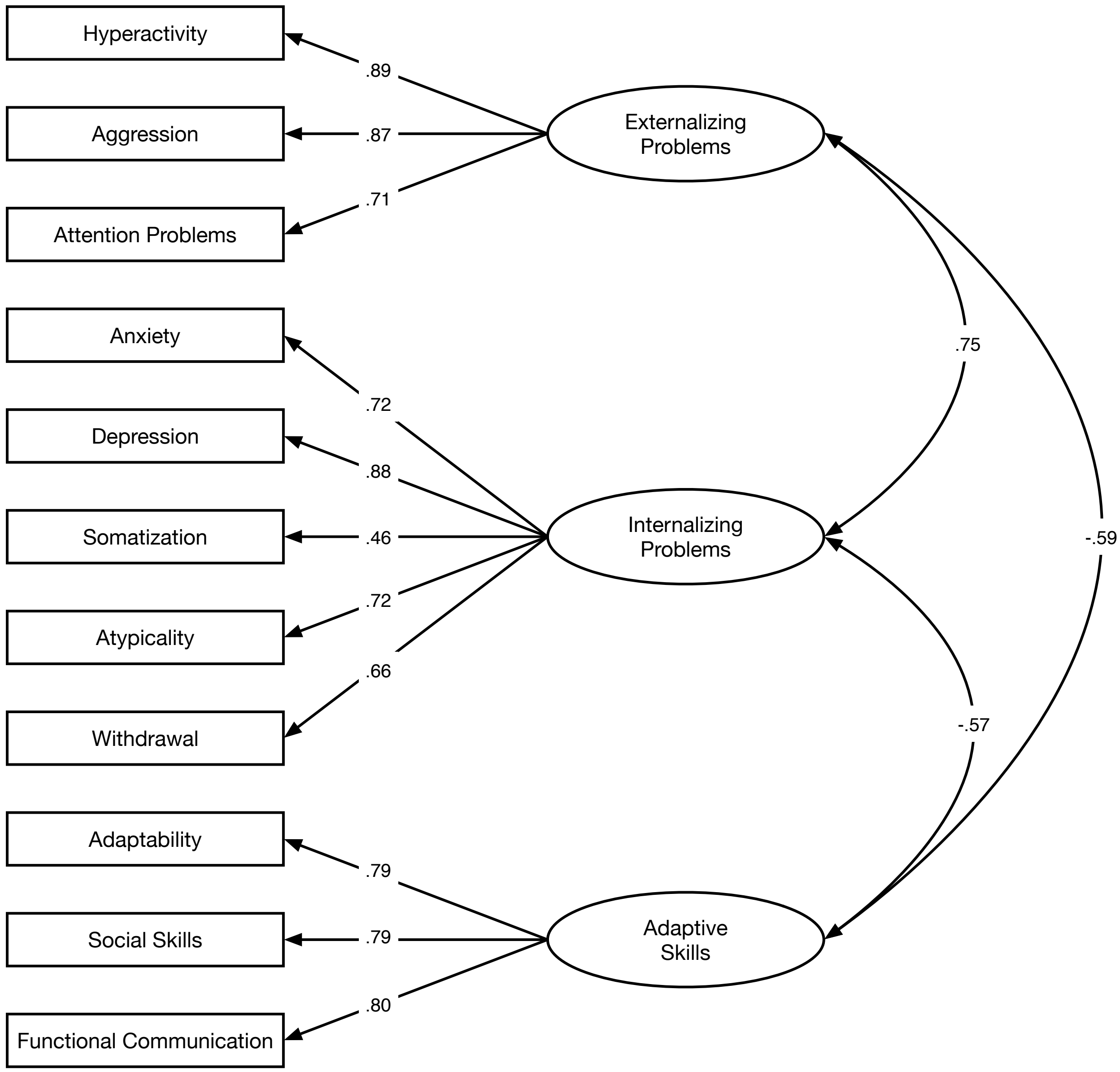

Figure 1. BASC-3 TRS Preschool Final Standardized Measurement Model Using Factor Loadings from BASC-3 *Manual* Table 9.17. $\chi^2(41) = 1{,}383.9$, CFI = .76, RMSEA = .20 Model fit statistics from BASC-3 *Manual* Table 9.16

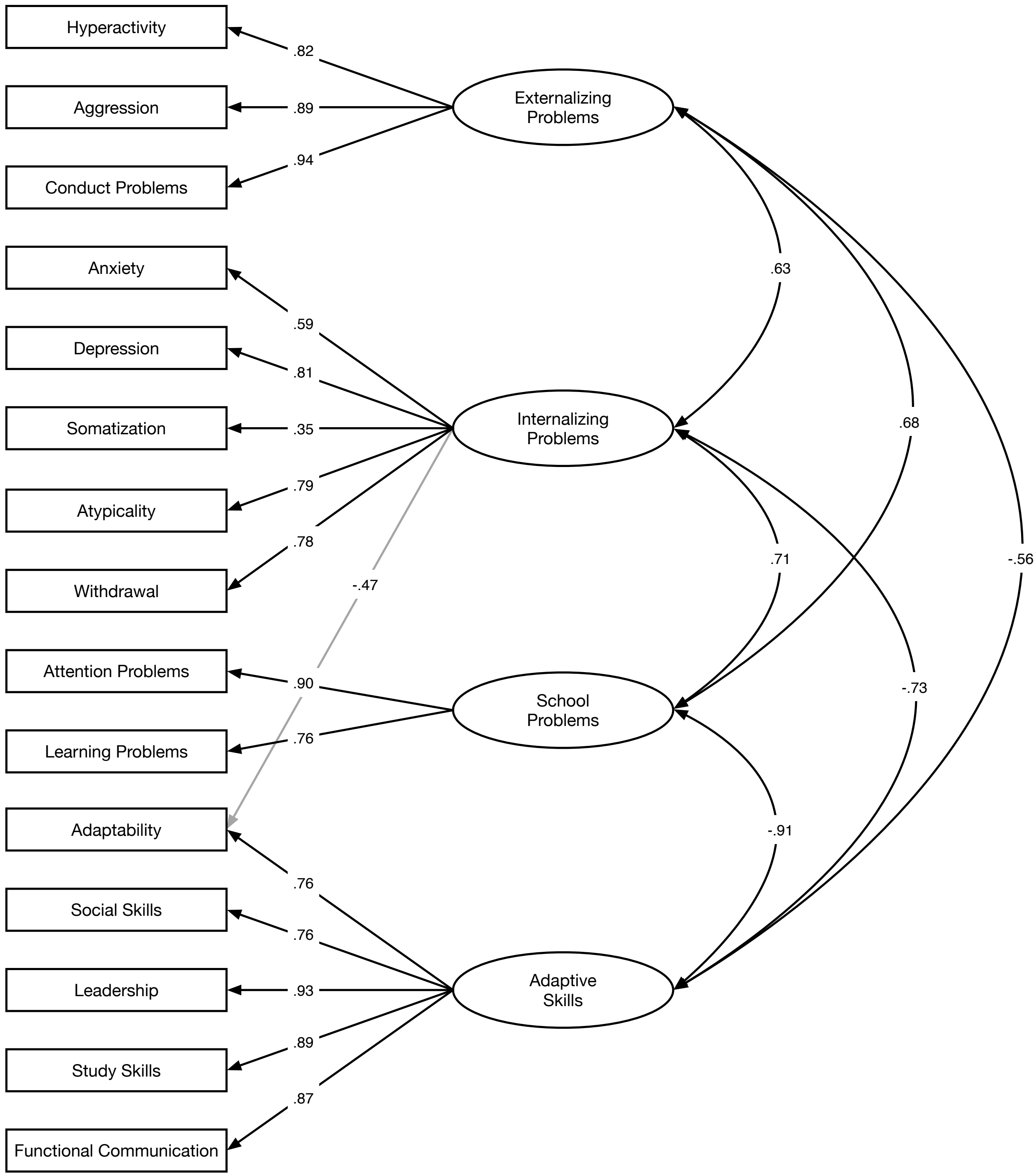

Figure 2. BASC-3 TRS Child Final Standardized Measurement Model Using Factor Loadings from BASC-3 *Manual* Table 9.17. $\chi^2(83) = 3{,}056.2$, CFI = .81, RMSEA = .18 Model fit statistics from BASC-3 *Manual* Table 9.16

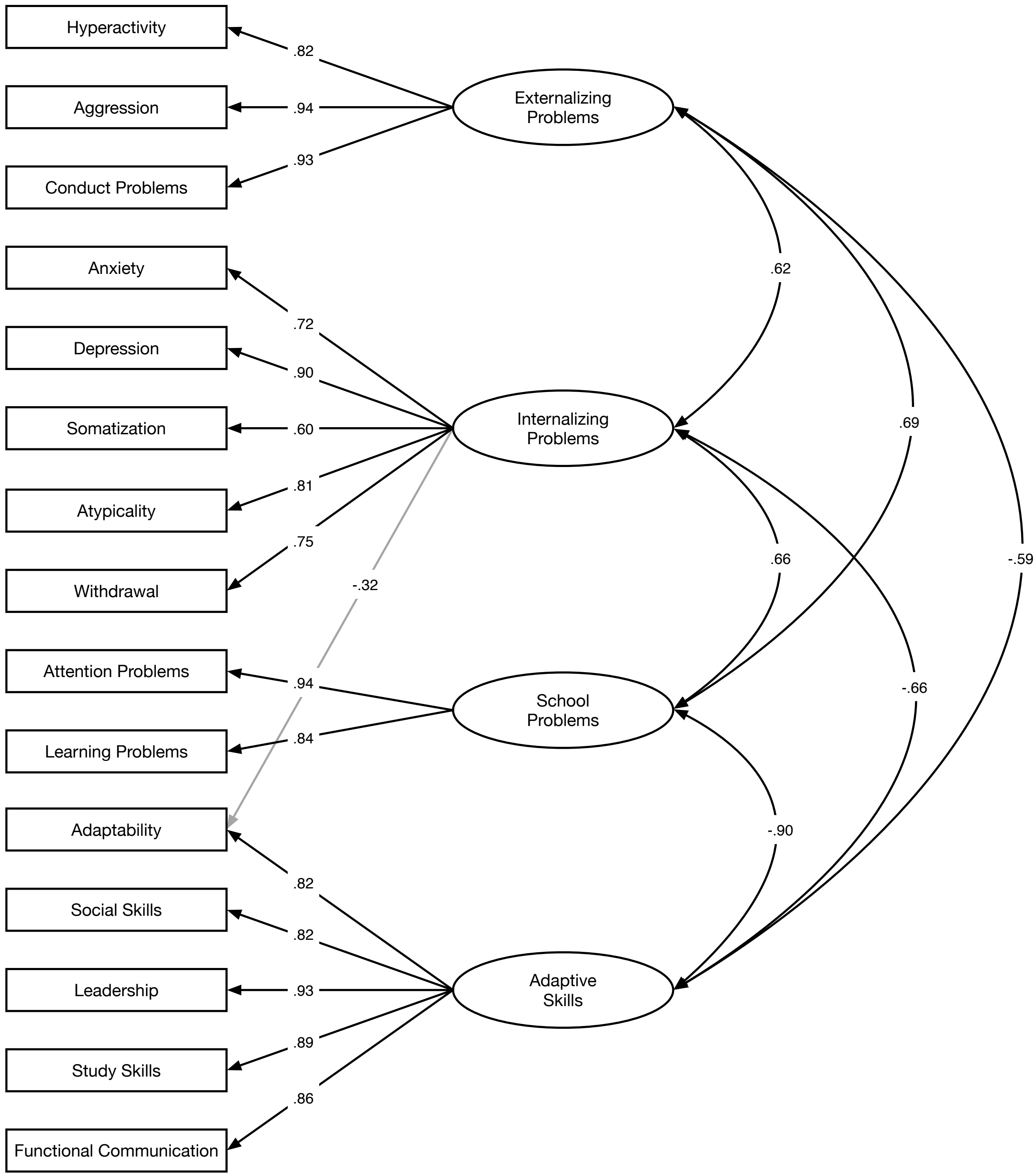

Figure 3. BASC-3 TRS Adolescent Final Standardized Measurement Model Using Factor Loadings from BASC-3 *Manual* Table 9.17. $\chi^2(83) = 2{,}632.4$, CFI = .82, RMSEA = .18 Model fit statistics from BASC-3 *Manual* Table 9.16