Article

Journal of Psychoeducational Assessment
2019, Vol. 37(2) 181–193

DOI: 10.1177/0734282917739109
journals.sagepub.com/home/jpa
SAGE

Examining the Latent Structure of the BASC-3 BESS Parent Preschool Form

Erin Dowdy[1], Christine DiStefano[2], Fred Greer[2], Stephanie Moore[1], and Kelvin Pompey[2]

Abstract

Screening for emotional and behavioral risk at the preschool level provides an opportunity to inform early intervention and prevention efforts. This study reports initial validation information for the *Behavioral Assessment System for Children–Third Edition, Behavioral and Emotional Screening System, Parent Form–Preschool* (BASC-3 BESS Parent-P). Using an Integrative Data Analysis framework, the BASC-3 BESS Parent-P latent structure was investigated using the norming sample from the BASC-3 (n = 459) as well as two randomly split samples from the BASC-2 norming sample (development sample n = 770; validation sample n = 799). Five models were tested using confirmatory factor analyses; findings suggested a four-factor oblique solution, including Internalizing Risk, Externalizing Risk, Adaptive Skills, and Attention Problems factors. Future research directions and use in school-based screening applications are presented.

Keywords

Behavioral and Emotional Screening System, screening, preschool, parent report

There is broad evidence supporting the value of early prevention and intervention of emotional and behavioral problems in school-based settings (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Although only approximately 12% of schools engage in a systematic screening process for emotional and behavioral risk (Bruhn, Woods-Groves, & Huggle, 2014), the practice is on the rise (Kamphaus, Reynolds, & Dever, 2014). The increased use of screening for behavioral and emotional risk (BER) is due to multiple reasons, including the emphasis on prevention and reliance on data-based decision making within widely adopted multitiered models of service delivery (e.g., multitiered systems of support, response to intervention), the provision of federal funds to support early identification and intervention efforts, the documented poor outcomes for students with emotional and behavioral problems, and the increased number of time and cost-efficient screening instruments (Bradley, Doolittle, & Bartolotta, 2008; Dowdy, Chin, & Quirk, 2013; U.S. Department of Education, 2006).

[1]University of California, Santa Barbara, CA, USA
[2]University of South Carolina, Columbia, SC, USA

Corresponding Author:
Erin Dowdy, Department of Counseling, Clinical, & School Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106-9490, USA.
Email: edowdy@ucsb.edu

Screening for emotional and behavioral risk may be especially critical at the preschool level (DiStefano & Kamphaus, 2007; Dowdy et al., 2013). Enrollment in prekindergarten has increased dramatically in the past decade, with estimates indicating that more than 1.3 million children (32% of all 3- and 4-year-olds) attend state-funded preschools (Barnett, Carolan, Squires, & Clarke Browne, 2013). Correspondingly, the number of children entering preschools with emerging social, behavioral, or emotional difficulties is also increasing. For example, among children ages 1 to 6 years, approximately 10% to 13% have emotional or behavioral disorders (Conroy & Brown, 2004). Furthermore, research suggests that behavioral and emotional problems that arise in early childhood are relatively stable and also predictive of negative educational and social outcomes (e.g., Lane, Little, Menzies, Lambert, & Wehby, 2010).

Prevention and early intervention services for social-emotional and behavioral problems have been recommended for preschoolers (Conroy & Brown, 2004), based on evidence supporting the positive outcomes following early intervention among young children (Brophy-Herb, Lee, Nievar, & Stollak, 2007). Although several methods are available for identifying children with behavioral or emotional risk (e.g., teacher nomination, pediatric referral, parent referral), school-wide universal screening that draws upon important informants for young children (i.e., parents and teachers) has been recommended as an optimal approach (Kamphaus et al., 2014).

School-wide universal screening can be accomplished at the preschool level by gathering information from parents or teachers. As informants, parents have been found to be ideal for providing information about their child's emotional and behavioral functioning (Smith, 2007). In addition, parents are able to provide information on their child's functioning earlier than teachers, as they need 4 to 6 weeks of familiarization with a child to accurately evaluate behavior. Also, parents may be eager to share information on their child's functioning, as preschool may be the first setting where parents have access to behavioral support services (DiStefano & Kamphaus, 2007).

Given benefits that may be realized if behavioral and emotional screening is conducted with preschool children, the focus of this psychometric investigation is on the newest edition of a commonly used screener: the *Behavioral Assessment System for Children–Third Edition, Behavioral and Emotional Screening System, Parent Form–Preschool* (BASC-3 BESS Parent-P; Kamphaus & Reynolds, 2015). The BASC-3 BESS is a broad screening instrument designed to identify behavioral and emotional strengths and weaknesses in children and adolescents. The BASC-3 BESS is a part of the latest edition of the BASC-3, which was updated from the BASC-2 BESS (Kamphaus & Reynolds, 2007) with improved normative data and with goals of improved item content, scale reliability, and score inference validity. In addition, due to studies using the BASC-2 BESS which supported a multifactorial interpretation of the items despite only one overall score being provided (e.g., Dever, Mays, Kamphaus, & Dowdy, 2012), one goal for the BASC-3 BESS was to identify subindex scores to aid in the interpretation of multiple areas of functioning (Kamphaus & Reynolds, 2015). To date, however, despite changes in item content and scores provided (i.e., the BASC-3 BESS provides an overall and subindex scores), rigorous investigation of the underlying structure of the most recent edition of the BESS has not yet been conducted. Consistent with the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), psychometric validation of the BASC-3 BESS Parent-P is needed.

To validate the structure, we used two samples: (a) the BASC-3 BESS Parent-P norming sample, published in 2015, and (b) the BASC-2 BESS Parent-P norming sample, published in 2007. As described later, the original BASC-2 BESS Parent-P and the revised BASC-3 BESS Parent-P versions differ; however, the same three constructs (Internalizing Risk, Externalizing Risk, and Adaptive Skills Risk) are measured and the operational definition of the constructs is the same. In addition, although the number of items differs, items assessing attention problems were included in both the BASC-2 (i.e., four items) and BASC-3 (i.e., two items) BESS Parent-P forms. The method of using measurement techniques to combine information from different scales measuring the same metric is a relatively new method suggested by measurement scientists, termed Integrative Data Analysis (IDA; Bauer & Hussong, 2009; Hussong, Curran, & Bauer, 2013). IDA has been recommended as a method to gain greater information from multiple scales or instruments that measure the same latent constructs. Thus, IDA is a novel method to examine and replicate the optimal structure of the BASC-3 BESS Parent-P, given that the three target constructs are identical. Therefore, the purpose of this investigation was to examine the latent structure of the BASC-3 BESS Parent-P using the BASC-3 norming sample and replicate findings with the BASC-2 norming sample.

Method

Instrumentation

To select the items for the published version of the BASC-3 BESS Parent-P, an initial item pool was created from the items retained following bias analyses on the BASC-3 standardization sample. Principal components analyses were performed on the items from each composite scale of the BASC-3 (e.g., Externalizing Risk, Internalizing Risk, and Adaptive Skills Risk) and a smaller pool of items was evaluated based on several criteria including content coverage and loading strength across forms (i.e., parent report, teacher report, student self-report) and levels (i.e., preschool, child/adolescent). Items with unique content, relatively high loadings, and similar psychometric properties across forms were further considered. Then, an iterative process of adding and removing items was employed with a focus on achieving adequate coverage of content across constructs and strong psychometric support.

The operational BASC-3 BESS Parent-P includes 29 items, with nine items assessing each of three broad behavioral dimensions. The Externalizing Risk dimension consists of items associated with externalizing behaviors, such as hyperactivity, aggression, and conduct problems (e.g., “Hits other children”). The Internalizing Risk dimension includes items assessing anxiety, depression, and somatization, which are characteristic of internalizing behaviors (e.g., “Is easily upset”). The Adaptive Skills Risk dimension assesses core characteristics of adaptive behavior including adaptability, social skills, and activities of daily living important for functioning at home and school, and in the community (e.g., “Responds appropriately when asked a question”). Two additional items assessing Attention Problems are also included, although these items do not belong to any specific dimension at the preschool level. The constructs are presumed to load onto a higher order factor, representing BER.

The BASC-2 BESS Parent-P form is similar in nature and structure to the BASC-3 BESS Parent-P. While the form includes 30 items instead of 29, more than half of the item content differs across forms; only 11 items from the previous version were retained. In addition, the number of items per dimension differs, with the BASC-2 form including nine items each measuring Externalizing and Internalizing Risk, eight items measuring Adaptive Skills Risk, and four measuring Attention Problems. The operational definitions of the constructs, however, are identical across the two BESS forms (Kamphaus & Reynolds, 2007, 2015). Per measurement theory assumptions, the selection of items on a given form may be thought of as a sampling of items from the available items in the empirical domain which measure the latent construct (e.g., Allen & Yen, 1979; Crocker & Algina, 1986). Given that confirmatory factor analysis (CFA) models examine the structure of the relations, the two norming samples may be used to assess the underlying theory.

When completing BASC-2 and BASC-3 BESS Parent-P forms, parents reflect on their child’s behaviors over the past 6 weeks and rate the frequency of behaviors. A 4-point response scale is provided for each item (i.e., 0 = *never*, 1 = *sometimes*, 2 = *often*, and 3 = *almost always*) and an overall score is created by summing item ratings. Items measuring Adaptive Skills are reverse-scored and then the total score is transformed to a *T* score (*M* = 50, *SD* = 10) to create the Behavioral and Emotional Risk Index (BERI), in which higher scores are indicative of higher levels of BER. Students are classified as having *Normal* (*T* scores of 60 or below; one standard deviation above the mean or lower), *Elevated* (*T* scores between 61 and 70; between one and two standard deviations above the mean), or *Extremely Elevated* (*T* scores of 71 or higher; more than two standard deviations above the mean) levels of risk. In addition to the BERI, three subindex scores are provided (i.e., Externalizing Risk Index, Internalizing Risk Index, and the Adaptive Skills Risk Index) with classification categories associated with raw scores (see Table 2.4 in the BASC-3 BESS manual for additional details; Kamphaus & Reynolds, 2015). Similar to the BERI, subindex classification ranges are based on standard deviations above the mean. As presented in the manual in more detail, the sensitivity, specificity, and positive and negative predictive power of these classification cut scores have been examined in relation to the BASC-3 full-length rating scale forms and known clinical groups (Kamphaus & Reynolds, 2015). The overarching goal was to develop cut scores that maximized the likelihood of identifying children with true behavioral and emotional problems (maximizing sensitivity and positive predictive power), while maintaining acceptable levels of specificity. In the Elevated risk score range, there are significantly more individuals who have been identified as having a behavioral or emotional problem (i.e., true positive), and the Extremely Elevated risk category indicates an even higher likelihood of a child having a behavioral or emotional problem. Within a screening context, it is advised that children who score in the Elevated or Extremely Elevated ranges are referred for additional assessment to further determine the possible presence of a problem.

Reliability and validity evidence for the BASC-3 BESS Parent-P is also presented in the manual. Estimates of internal consistency include split-half estimates for the BERI (.95-.96), and coefficient alpha estimates for the three subindex scores of internal consistency, ranging from .85 for the Internalizing Risk to .88 for Externalizing Risk. Three-week test–retest reliability estimates range from .81 to .89, and interrater reliability estimates range from .63 to .74. Validity evidence is presented in the form of comparisons of the BASC-3 BESS Parent-P scores with scores from other tests of similar constructs including the full BASC-3 Parent Rating Scales (Kamphaus & Reynolds, 2015) and the Achenbach System of Empirically Based Assessment (Achenbach & Rescorla, 2001).

In the manual, CFA was employed to evaluate the underlying model. Specifically, Externalizing Risk, Internalizing Risk, and Adaptive Skills Risk were included as first-order factors, and BERI was included as a second-order factor. The results indicated “marginal” model fit for the BESS Parent-P, although specific fit indices were not provided (Kamphaus & Reynolds, 2015).

Participants

Standardization data collection for the BASC-3 occurred from April 2013 to November 2014 with the goal of obtaining data from a large and representative sample of children in the United States. The BASC-3 sample included ratings of 459 BESS Parent-P forms, collected as part of the norming sample. The sample of parents was obtained from across the United States, with the majority from the south. Parents varied in their level of education, with most (33%) reporting between 1 and 3 years post–high school education. Of the children rated, the average age was 4.5 years (*SD* = 10.3 months), with approximately 49% males and 51% females. The sample was diverse in regard to race/ethnicity, with slightly more than half of the children rated as Caucasian (51.2%). The majority of children (approximately 93%) did not have a prior clinical diagnosis (e.g., autism, attention deficit hyperactivity disorder).

As the norming sample from the BASC-3 BESS was too small to split into subsamples, the BASC-2 BESS norming dataset was used for validation. Although the two BESS screeners are not the same, consistent with IDA methodology, the same constructs are measured and the constructs share the same operational definition. The norming sample for the BASC-2 BESS Parent-P forms consisted of 1,835 ratings. For validation purposes, this sample was randomly split into a development and a validation sample. The development sample consisted of 918 children, and the validation sample consisted of 917 children. As the BASC-2 BESS Parent-P norming sample included participants outside of the 3 to 5 years, cases outside of this age range were not included in analyses. The developmental sample for analyses consisted of 770 ratings comprised predominantly of children identified as Caucasian (66%) and male (54%) with a mean age of 4.4 years (*SD* = 9.5 months). The final validation sample consisted of 799 ratings; the majority of children were Caucasian (64%) and male (52%) with a mean age of 4.4 years (*SD* = 9.6 months).

Statistical Methods

Analyses were conducted using the M*plus* software program (Version 7.4; Muthén & Muthén, 1998-2015) using weighted least squares with mean and variance adjustment (WLSMV) estimation. This technique has been recommended for analysis of categorical variables (Finney & DiStefano, 2013). WLSMV adjusts the chi-square fit statistic, providing a value that more closely approximates what would be achieved if normal data were used (Finney & DiStefano, 2013). Also, standard errors of parameter estimates are adjusted with the robust method, producing estimates that more closely approximate true values (Muthén & Muthén, 1998-2015). In addition, the pattern-matching technique was used to accommodate profiles with missing responses. This is the default option to deal with missing data when WLSMV estimation is used.

CFA was used as the BASC-3 BESS Parent-P was developed from a strong theoretical perspective. The tested models were developed from a combination of prior BESS research, the hypothesized theoretical structure of the screener, and recommendations for testing multifactor models (e.g., Gignac, 2007). Attention Problems items are not included as part of the three dimensions (i.e., Externalizing Risk, Internalizing Risk, and Adaptive Skills Risk) at the preschool level and are generally not included in validation analyses (e.g., Kamphaus & Reynolds, 2007, 2015). However, as noted in the manuals, these items do contribute to a total BER Index score and previous analyses with the BASC-2 with the Teacher Rating Scale-Preschool form showed that the Attention Problem items were important to include (DiStefano, Greer, & Kamphaus, 2013). Thus, models were tested that both did and did not include Attention Problem items.

To identify the optimal structure underlying the BASC-3 BESS Parent-P, four different multifactor models were tested. The CFA models tested, replicated, and extended the higher order investigations presented in the manual: (a) a higher order model where three first-order latent dimensions (i.e., Externalizing Risk, Internalizing Risk, and Adaptive Skills Risk) are related to an overarching (i.e., second order) BERI and (b) a higher order model where four first-order latent dimensions (i.e., Externalizing Risk, Internalizing Risk, and Adaptive Skills Risk, and Attention Problems) are related to an overarching BERI. In addition, multifactor models were tested including (c) a three-factor oblique model where items loaded on their hypothesized dimension and (d) a four-factor oblique model, including Attention Problems, where items loaded on their hypothesized dimension.

Models were evaluated using a selection of four fit indices, where selected indices focus on different aspects of model fit. Fit indices were chosen on the basis of recommendations from previous research (e.g., Hu & Bentler, 1998; Tanaka, 1993): (a) chi-square statistic, (b) nonnormed fit index (NNFI or Tucker–Lewis index [TLI]), (c) comparative fit index (CFI), and (d) root mean square error of approximation (RMSEA). All fit indices are included with the program output. Both the NNFI/TLI and CFI are incremental fit indices and test the proportionate improvement in fit by comparing the target model with a baseline model with no correlations among observed variables (Hu & Bentler, 1998). NNFI and CFI values approximating at least .95 were indicative of good fit (Hu & Bentler, 1998) and .90 or above for acceptable fit. The RMSEA represents closeness of fit between the models and should approximate .05 to demonstrate close fit of the model (Browne & Cudeck, 1993). The 90% confidence interval (CI) around the RMSEA point estimate should contain .05 to indicate the possibility of close fit; values of .08 or below indicate acceptable fit (Browne & Cudeck, 1993). When WLSMV is used with ordinal data, fit indices that use the chi-square in their calculations are also corrected and may be considered robust fit indices (Finney & DiStefano, 2013). In addition, completely standardized parameter estimates of indicators and modification indices were evaluated for the optimal model. Also, relations between latent variables and percentage of variance shared between items and constructs were examined.

Results

BASC-3 Norming Sample

Fit indices for each of the tested BASC-3 BESS Parent-P models are presented in Table 1. In testing the higher order Models A and B, a Heywood case (Dillon, Kumar, & Mulani, 1987) was observed; a loading greater than 1 and negative residual variance were evident for the Externalizing Risk factor relating to the higher order factor. Model fit statistics for the corrected model, fixing the residual variance for the Externalizing Risk factor to 0 (Dillon et al., 1987), exhibited adequate fit (see Table 1).

Given the errors that occurred when specifying both higher order models, three- (Model C) and four- (Model D) factor oblique solutions were examined. Each of Model C and Model D achieved adequate fit. Overlap in the RMSEA CIs indicated equivalence of fit for the two models. Model E, a one-factor unidimensional model, was also tested and exhibited poor fit to the data (Table 1). Given that the BASC-3 BESS Parent-P was designed to include two Attention Problems items, Model D was selected for further consideration.

BASC-2 Norming Samples

The same series of models that were tested with the BASC-3 BESS Parent-P norming dataset were tested with the BASC-2 BESS Parent-P norm samples. The indices are included in Table 1. Across the set of models, the fit indices do not meet generally accepted guidelines for good fit. Fit for both the development and validation samples were similar, and are discussed together.

As with the BASC-3 BESS Parent-P norm sample, the one-factor model (Model E) did not fit the data well; this model was not considered further. Different from the previous analyses, the higher order models (Models A and B) did not show estimation problems; however, there are four Attention Problems included on the BASC-2 BESS Parent-P which may have resulted in greater stability with estimation. While both the higher order models fit acceptably, the multifactor models were an improvement. Considering both the three- (Model C) and four- (Model D) factor designs, the four-factor model yielded optimal fit. This model had the lowest chi-square value, highest values of CFI and TLI, and an RMSEA value suggesting acceptable fit; however, these fit indices were comparable with the other models. As such, we chose Model D due to a mix of both conceptual and statistical information. Overall, as with the BASC-3 BESS Parent-P norm sample, the four-factor CFA model (Model D) was

Table 1. Fit Indices From BESS Analyses.

Model	χ^2	*df*	CFI	TLI	RMSEA (90% CI)
BASC-3 norm sample (*n* = 459)					
Three-factor higher order (A[a])	1,081.81***	322	.921	.914	.072 [.067, .076]
Four-factor higher order (B[a])	1,191.46***	374	.922	.915	.069 [.065, .073]
Three-factor oblique (C)	1,111.22***	321	.918	.911	.073 [.069, .078]
Four-factor oblique (D)	1,234.37***	371	.918	.910	.071 [.067, .076]
Unidimensional (E)	2,681.53***	377	.780	.763	.115 [.111, .120]
BASC-2 norm development sample (*n* = 770)					
Three-factor higher order (A)	2,555.83***	402	.881	.871	.083 [.080, .087]
Four-factor higher order (B)	2,895.61***	401	.862	.850	.090 [.087, .093]
Three-factor oblique (C)	2,555.83***	402	.881	.871	.083 [.080, .087]
Four-factor oblique (D)	2,415.47***	399	.889	.878	.081 [.078, .084]
Unidimensional (E)	3,808.53***	405	.812	.798	.104 [.101, .108]
BASC-2 norm validation sample (*n* = 779)					
Three-factor higher order (A)	2,701.96***	402	.891	.882	.085 [.082, .088]
Four-factor higher order (B)	3,005.81***	401	.877	.867	.090 [.087, .093]
Three-factor oblique (C)	2,701.96***	402	.891	.882	.085 [.082, .088]
Four-factor oblique (D)	2,592.29***	399	.896	.887	.083 [.080, .086]
Unidimensional (E)	4,113.99***	405	.825	.812	.107 [.104, .110]

Note. BESS = Behavioral and Emotional Screening System; χ^2 = chi-square test of model fit; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; BASC = Behavioral Assessment System for Children.
[a]Residual variance for the Externalizing Risk factor was constrained to 0 to allow for model convergence.
***$p < .001$.

chosen as the best-fitting model for both samples tested with the BASC-2 BESS Parent-P norming datasets.

Comparing across the two norming samples, the BASC-3 BESS Parent-P four-factor CFA model yielded better fit than those tested with the BASC-2 samples. With the BASC-2 norm sample, the RMSEA value approximated good fit, whereas the model fit was at the acceptable level with the BASC-3 norm sample. In addition, TLI/CFI values were also higher, and chi-square model fit values were lower.

Model Selection Across Samples

Standardized loadings, factor correlations, and reliability of factor measurement coefficients for the optimal four-factor oblique solution, Model D, for each of the BASC-3 BESS Parent-P and BASC-2 BESS Parent-P development and validation samples are presented in Table 2.

Across all samples, relations between items and their associated factors were generally strong. Interfactor correlations among each factor were strong across samples, with estimates ranging from $r = .46$ to $r = .91$. In addition, examination of model fit indices across samples did not offer support for theoretically indicated modifications to the factor structure of Model D. Composite reliability and average variance extracted estimates illustrated strong reliability of each factor across samples (Hancock & Mueller, 2001; Raykov, 2004). Notably, the average variance extracted for the Internalizing Risk Index is stronger in the BASC-3 BESS than in the BASC-2 BESS. Therefore, the four-factor oblique solution, including each of an Internalizing Risk Index, Externalizing Risk Index, Adaptive Skills Index, and Attention Problems Index factor, was selected as the final, best-fitting and theoretically indicated model.

Table 2. Standardized Factor Loadings, Factor Correlations, and Reliability of Model D Across Samples.

	BASC-3 norm sample				BASC-2 norm sample BESS development (validation)			
Item	EXT	INT	ADAP	AP	EXT	INT	ADAP	AP
Is unable to slow down.	.71							
Is overly aggressive.	.87							
Needs too much supervision.	.76							
Argues when denied own way.	.66							
Disrupts the play of other children.					.66 (.69)			
Hits other children.	.65				.71 (.62)			
Has poor self-control.					.80 (.70)			
Annoys others on purpose.					.63 (.55)			
Defies people in authority.	.65				.62 (.69)			
Throws tantrums.	.73				.70 (.73)			
Loses temper too easily.	.87				.72 (.66)			
Acts without thinking.					.71 (.66)			
Acts out of control.	.88				.87 (.82)			
Quickly loses interest in things.						.74 (.72)		
Whines.						.50 (.54)		
Changes moods quickly.						.59 (.62)		
Has trouble falling asleep.						.57 (.52)		
Is happy.						.58 (.52)		
Is easily frustrated.		.55				.75 (.72)		
Gets very upset when things are lost.						.42 (.38)		
Has trouble sleeping through the night.						.58 (.47)		
Gets upset when away from favorite toy or object.						.37 (.37)		
Complains of pain.		.59						
Is easily upset.		.79						
Is nervous.		.56						
Says "Nobody likes me."		.55						
Is negative about things.		.80						
Complains of physical problems.		.63						
Worries about things that cannot be changed.		.56						
Is irritable.		.82						
Is a "good sport."								
Adjusts well to changes in routine.			.70					
Responds appropriately when asked a question.			.81					
Adjusts easily to new surroundings.			.59					
Is easily calmed when angry.			.78					
Gets along well with others.			.81					
Politely asks for help.			.74				.77 (.75)	

(continued)

Table 2. (continued)

	BASC-3 norm sample				BASC-2 norm sample BESS development (validation)			
Item	EXT	INT	ADAP	AP	EXT	INT	ADAP	AP
Begins conversations appropriately.			.68				.65 (.65)	
Communicates clearly.			.76				.77 (.76)	
Is able to describe feelings accurately.			.68				.72 (.75)	
Ignores safety rules.							.70 (.70)	
Tries new things.							.55 (.67)	
Needs help putting on clothes.							.46 (.40)	
Pays attention.								.87 (.89)
Has a short attention span.				.78				.80 (.78)
Listens carefully.								.87 (.86)
Listens to directions.								.83 (.81)
Has trouble concentrating.				.87				
	Factor intercorrelation matrices							
Factors	EXT	INT	ADAP	AP	EXT	INT	ADAP	AP
EXT	1				1			
INT	.83	1			.86 (.91)	1		
ADAP	.57	.46	1		.61 (.58)	.66 (.66)	1	
AP	.77	.71	.47	1	.69 (.71)	.74 (.76)	.87 (.86)	1
	EXT	INT	ADAP	AP	EXT	INT	ADAP	AP
Composite reliability	.92	.89	.91	.81	.90 (.89)	.81 (.79)	.91 (.90)	.86 (.87)
Average variance extracted	.57	.49	.53	.69	.51 (.47)	.34 (.31)	.71 (.70)	.45 (.47)

Note. All item loadings and factor intercorrelations are significant at $p < .001$. Bolded items represent items that were present in both versions of the BASC. BASC = Behavioral Assessment System for Children; BESS = Behavioral and Emotional Screening System; EXT = externalizing problems; INT = internalizing problems; ADAP = adaptive skills; AP = attention problems.

Discussion

This study sought to examine the latent structure of the BASC-3 BESS Parent-P. Through the use of IDA methodology, data from multiple instruments that measure the same latent construct of BER were combined to examine the latent structure of the BASC-3 BESS Parent-P. Specifically, three samples including the BASC-3 norming sample, and development and validation subsamples from the BASC-2 norming sample, were used to systematically investigate the underlying factor structure among U.S. preschool children. Based on previous research, the hypothesized structure of the screener, and recommendations for testing multifactor models, five models were tested via CFA.

In comparing the results from the BASC-3 and BASC-2 Parent-P samples, as well as comparing models within samples, the differences in fit indices were small; however, across the three samples, the four-factor oblique model yielded slightly better fit over the other models tested and was also theoretically supported. Therefore, this model was selected as the optimal model. This model consists of the following four factors: Internalizing Risk Index, Externalizing Risk Index, Adaptive Skills Index, and Attention Problems. This factor structure differs somewhat from the factor structure suggested in the manual for the BASC-3 BESS. Specifically, the factor structure provided in the manual indicates the presence of Internalizing, Externalizing, and Adaptive Skills factors. However, although two items designed to assess attention problems are a part of the screening form and are included in an overall BERI score, there is not an Attention Problems factor or subscale score provided. A similar structure was found to be optimal with previous analyses using the BASC-2 BESS Teacher-Preschool version (DiStefano et al., 2013).

The overall BERI score is calculated when scoring the published version of the BASC-3 Parent-P; however, results from this study indicate that the higher order models with an overarching second-order BERI were not supported with the BASC-3 BESS sample. Thus, the findings suggest that the BASC-3 BESS Parent-P is not adequately capturing an overarching BERI. As higher order models provided acceptable solutions with the BASC-2 BESS norm dataset, it is not clear if the results are sample specific. Further study may examine larger samples of parent ratings using the BESS-3 instrument as well as validity studies investigating subscale scores and a total BERI to determine if a higher order model is viable.

Results did show support for an Attention Problems factor. As Attention Problems is a separate factor with the child scales, there may be support for including this factor in preschool samples as well. Considering that attention is a key temperamental variable related to a variety of educational and life outcomes (Molina & Pelham, 2003) and that prior research with the teacher-rated version of the BASC-2 preschool form highlighted the importance of the attention items (DiStefano et al., 2013), practitioners and researchers may wish to consider an Attention Problems factor to identify students who may benefit from behavioral interventions designed to mitigate attention problems (see Vannest, Reynolds, & Kamphaus, 2015, for intervention suggestions).

In addition, practitioners should be cautioned about an overreliance on, and interpretation of, the overarching BERI, as current results do not support its use. Although we considered testing a bifactor model with an overall BER factor and specific factors of Externalizing, Internalizing, Adaptive Skills, and Attention, we chose not to engage in a model-fitting comparison between bifactor and other higher order models due to the inherent statistical bias favoring the fit of the bifactor models (Murray & Johnson, 2013; Rodriguez, Reise, & Haviland, 2016). Were the BER theoretical structure more strongly supported, the BERI score could serve as a general “red flag” to trigger additional assessment (Dowdy, Dever, Raines, & Moffa, 2016). In multiple gating screening, a broad first “gate” screening is offered universally, and those identified to be at risk are then referred for an additional gate of more intensive and precise assessments. The BASC-3 manual suggests the overall BERI score be used as a first gate score, and that those receiving an Elevated *T* score may require the longer, omnibus BASC-3 Rating Scale (Kamphaus & Reynolds, 2015). Considering results of this study, practitioners may instead wish to use the Externalizing Risk Index, Internalizing Risk Index, Adaptive Skills Index, and Attention Problems items as a first gate score. Given the overarching goals of a first gate screening to identify any student who may benefit from additional assessment, the use of four index scores as opposed to one overall BER Index or three subindex scores may actually serve to identify more students in need. Future research to determine suggested guidelines or cut scores for the use of the four index scores will be helpful; however, prior to these guidelines being established, practitioners can consider scores that are in the Elevated or Extremely Elevated range (*T* scores of 61 and above) as triggering the need for additional assessment. Although additional research is needed to investigate the value of multiple gates, balancing both sensitivity and specificity (Stiffler & Dever, 2015), practitioners using the BASC-3 BESS Parent-P are advised not to rely on the BERI score.

This study incurred several limitations which may also provide direction for future research. First, although it provided for the use of a novel methodology (IDA), the use of the BASC-2 BESS Parent-P norming sample to test the latent structure of the BASC-3 BESS Parent-P form was not ideal. Specifically, the exact items were not present across all samples and the overlapping items were presented in a different manner to the various sample participants. Further research to clarify the latent structure of the BASC-3 BESS Parent-P among U.S. preschoolers is warranted to replicate the findings with additional samples. Second, this examination of structural validity only provides information on the latent structure of the form and does not investigate other important psychometric areas, including measurement and structural invariance across important demographic factors or the ability to predict later educational outcomes. Although some external validity estimates are provided in the manual (Kamphaus & Reynolds, 2015), additional independent research is needed. Finally, we acknowledge that there are other BASC-3 screening forms for use with preschoolers and older children across multiple raters that require additional psychometric analyses.

Overall, this study seeks to provide continued support for the use of school-based screening in early childhood using parents as informant by providing initial psychometric evidence in support of the BASC-3 BESS Parent-P. This practice of gathering parent screening data to inform further assessment and intervention is aligned with early childhood standards (e.g., National Association for the Education of Young Children, Division of Early Childhood) that stress the importance of acquiring parental input when assessing preschool children. As psychometrically sound screening tools provide the foundation of assessment and intervention practices, it will continue to be important to advance the science of screening through the further investigation of screening measures used in schools.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education.

Acknowledgment

The authors gratefully acknowledge Pearson for granting use of the Behavioral Assessment System for Children, Parent Rating Scales–Preschool (BASC PRS-P) norming dataset.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant #R305A150152 to South Carolina Research Foundation.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA preschool forms & profiles*. Burlington: Research Center for Children, Youth, & Families, The University of Vermont.

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Barnett, S. W., Carolan, M. E., Squires, J. H., & Clarke Brown, K. (2013). *The state of preschool 2013*. New Brunswick, NJ: National Institute for Early Education Research.

Bauer, D. J., & Hussong, A. M. (2009). Psychometric approaches for developing commensurate measures across independent studies: Traditional and new models. *Psychological Methods*, *14*, 101-125. doi:10.1037/a0015583

Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education*, *17*, 4-23. doi:10.1007/s10864-007-9058-6

Brophy-Herb, H., Lee, R. E., Nievar, M. A., & Stollak, G. (2007). Preschoolers' social competence: Relations to family characteristics, teacher behaviors and classroom climate. *Journal of Applied Developmental Psychology*, *28*, 134-148. doi:10.1016/j.appdev.2006.12.004

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bruhn, A. L., Woods-Groves, S., & Huggle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education and Treatment of Children*, *37*, 611-634. doi:10.1353/etc.2014.0039

Conroy, M. A., & Brown, W. H. (2004). Early identification, prevention, and early intervention with young children at risk for emotional or behavioral disorders: Issues, trends, and a call for action. *Behavioral Disorders*, *29*, 224-236.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt, Rinehart and Winston.

Dever, B., Mays, K., Kamphaus, R. W., & Dowdy, E. (2012). The factor structure of the BASC-2 Behavioral and Emotional Screening System Teacher Form, child/adolescent. *Journal of Psychoeducational Assessment*, *30*, 488-495. doi:10.1177/0734282912438869

Dillon, W. R., Kumar, A., & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin*, *101*, 126-135.

DiStefano, C. A., Greer, F. W., & Kamphaus, R. W. (2013). Multifactor modeling of emotional and behavioral risk of preschool-age children. *Psychological Assessment*, *25*, 467-476. doi:10.1037/a0031393

DiStefano, C. A., & Kamphaus, R. W. (2007). Development and validation of a behavioral screener for preschool-age children. *Journal of Emotional and Behavioral Disorders*, *15*, 93-102. doi:10.1177/10634266070150020401

Dowdy, E., Chin, J., & Quirk, M. (2013). Preschool screening: An examination of the Behavioral and Emotional Screening System Preschool Teacher Form (BESS preschool). *Journal of Psychoeducational Assessment*, *31*, 578-584. doi:10.1177/0734282913475779

Dowdy, E., Dever, B. V., Raines, T., & Moffa, K. (2016). A preliminary investigation into the added value of multiple gates and informants in universal screening for behavioral and emotional risk. *Journal of Applied School Psychology*, *32*, 178-198. doi:10.1080/15377903.2016.1165327

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405-432. doi:10.1111/j.1467-8624.2010.01564.x

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed.) (pp. 439–492). Charlotte, NC: Information Age Publishing.

Gignac, G. E. (2007). Multi-factor modeling in individual differences research: Some recommendations and suggestions. *Personality and Individual Differences*, *42*, 37-48. doi:10.1016/j.paid.2006.06.019

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

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424-453. doi:10.1037/1082-989X.3.4.424

Hussong, A. M., Curran, P. J., & Bauer, D. J. (2013). Integrative data analysis in clinical psychology research. *Annual Review of Clinical Psychology*, *9*, 61-89. doi:10.1146/annurev-clinpsy-050212-185522

Kamphaus, R. W., & Reynolds, C. R. (2007). *Behavior Assessment System for Children–Second Edition (BASC-2): Behavioral and Emotional Screening System (BESS)*. Bloomington, MN: Pearson.

Kamphaus, R. W., & Reynolds, C. R. (2015). *Behavior Assessment System for Children—Third Edition (BASC-3): Behavioral and Emotional Screening System (BESS)*. Bloomington, MN: Pearson.

Kamphaus, R. W., Reynolds, C. R., & Dever, B. V. (2014). Behavioral and mental health screening. In R. J. Kettler, T. A. Glover, C. A. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational* *settings: Evidence-based decision making for schools* (pp. 249-273). Washington, DC: American Psychological Association.

Lane, K. L., Little, A. L., Menzies, H. M., Lambert, W., & Wehby, J. H. (2010). A comparison of students with behavioral challenges educated in suburban and rural settings: Academic, Social, and behavioral outcomes. *Journal of Emotional and Behavioral Disorders*, *18*, 131-148.

Molina, B. S., & Pelham, W. E. (2003). Childhood predictors of adolescent substance abuse in a longitudinal study of children with ADHD. *Journal of Abnormal Psychology*, *112*, 497-508.

Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, *41*, 407-422.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy*, *35*, 299-331.

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

Smith, S. R. (2007). Making sense of multiple informants in child and adolescent psychopathology: Guide for clinicians. *Journal of Psychoeducational Assessment*, *25*, 139-149. doi:10.1177/073428906296233

Stiffler, M. C., & Dever, B. V. (2015). *Mental health screening at school: Instrumentation, implementation, and critical issues*. New York, NY: Springer.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10-40). Newbury Park, CA: Sage.

U.S. Department of Education. (2006). *IDEA regulations: Early intervening services*. Washington, DC: Office of Special Education Programs. Available from http://idea.ed.gov/

Vannest, K. J., Reynolds, C. R., & Kamphaus, R. W. (2015). *BASC-3: Intervention guide*. Bloomington, MN: Pearson.