



Confirmatory Factor Structure and Predictive Validity of the Early Identification System—Student Report in a Community Sample of High School Students

Aaron M. Thompson¹ · Francis Huang¹ · Tyler Smith¹ · Wendy M. Reinke¹ · Keith C. Herman¹ · The County Schools Mental Health Coalition¹

Accepted: 12 October 2020 / Published online: 18 November 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

The purpose of this paper is to confirm the factor structure, examine the invariance, and investigate the predictive validity using disciplinary data for 5262 high school students who completed the Early Identification System—Student Response (EIS-SR). The development and theory of the EIS-SR is discussed along with prior work. Building off of prior factor analytic work with a separate sample, it was hypothesized the items of the EIS-SR would coalesce into seven factors representing Externalizing Behavior, Internalizing Behavior, Peer Relationship Problems, School Disengagement, Emotional Dysregulation, Attention and Academic Issues, and Relational Aggression. Furthermore, it was hypothesized that EIS-SR scores would be invariant with regard to gender and grade level. Lastly, it was hypothesized that students with high EIS-SR subscale scores would be predictive of school discipline events. Our analyses indicated the EIS-SR did fit the previously observed factor structure with the items loading on seven distinct scales. Tests for measurement invariance indicated support that the EIS-SR measured the seven factors equally well regardless of both gender and grade level. Lastly, EIS-SR subscale scores predicted spring office disciplinary referrals, both in and out of school suspensions, and attendance.

Keywords Confirmatory factor analysis · School mental health · Measurement · Screening · Predictive validity

Introduction

Universal screening has long been, currently is, and will remain a common method of identification, monitoring, and directing educational practices and programming for students in American public schools. Educational professionals screen students to get a better understanding of academic strengths and weaknesses (Gersten et al., 2012; January, Ardoin, Christ, Eckert, & White, 2016; Saddler & Asaro-Saddler, 2013). Students are screened for early signs of health-related concerns as mandated in most states for hearing and vision (Boesen & Lykke, 2012), and other states require screenings for exposure to lead and environmental toxins (CEH, 2005), scoliosis (Kadhim et al., 2020), and dental concerns (Arora et al., 2019). Students entering

school are also screened for developmental delays or early markers of disabilities (Lipkin & Macias, 2020). These screening data are used to identify struggling students and inform practices and programming to support those students.

Compared to the variety of regular screening practices in schools, efforts to screen students for social, emotional, and behavioral health needs are noticeably lagging. In 2005, it was estimated that 2% of public schools in the USA were administering regular screenings for mental health-related concerns (Romer & McIntosh, 2005), a practice that increased to 12.6% in 2010 (Bruhn, Woods-Groves, & Huddle, 2014) and 15.5% of schools by 2015 (Siceloff, Bradley, & Flory, 2017)—a trend that might equate to roughly 20% of schools engaging in mental health screenings by 2021. The glacial pace of mental health screening practices in school settings is concerning when contrasted alongside the increasing severity and prevalence of youth mental health struggles. For example, emergency rooms across the USA have reported a spike over the past decade in pediatric psychiatry cases requiring hospitalization (Chedekel, 2017). Teen suicide rates are

✉ Aaron M. Thompson
thompsonaa@missouri.edu

¹ School of Social Work, University of Missouri, 718 Clark Hall, Columbia, MO 65211, USA

at an all-time high with the Centers of Disease Control reporting a 57% increase between 2007 and 2017 (Curtin & Herron, 2019). The increasing rates of suicide over the past decade are paralleled by spikes in self-reported anxiety and depression (Ducharme, 2018) with 1 in 5 school-aged youth exhibiting symptoms of a mental health disorder to a degree that daily functioning is adversely impacted (Anderson & Cardoza, 2016; Merikangas et al., 2010). More alarming, studies suggest 80% of those youth will not access effective mental health services (Merikangas et al., 2010). In addition to these worsening indicators, the stress of the COVID-19 health crisis has required many students to attend school remotely, leaving teachers without daily face-to-face contact with struggling students. This might be enough to make mental health screening of our youth in schools a priority. If youth are not screened properly, social, emotional, and behavioral health needs are likely to remain unaddressed (Merikangas et al., 2010; Power, Eiraldi, Clarke, Mazzuca, & Krain, 2005), leading to significant challenges for educational professionals. For instance, unattended mental health problems contribute to behavioral disruptions which contribute to declines in teacher self-efficacy, emotional exhaustion, and burnout (Skaalvik & Skaalvik, 2007; Smith, 2019). Students who struggle with unattended mental health begin experiencing a cascade of behavioral health risks including school refusal and chronic absenteeism (Maynard et al., 2018; Stempel, Cox-Martin, Bronsert, Dickenson, & Allison, 2017), low achievement and disengagement (Battaglia et al., 2017), and eventually school dropout (Holen, Waaktaar, & Sagatun, 2018).

Fortunately, schools can do something about the problem. Randomized research studies conducted in natural school settings have repeatedly revealed that carefully implemented, data-driven, tiered models of prevention not only mitigate proximal social, emotional, and behavioral health stressors for 20% of youth experience on a daily basis (Arora et al., 2019; Barnett, Eber & Weist, 2013; Bradshaw, Reinke, Brown, Bevans, & Leaf, 2008; McIntosh, Ty, & Miller, 2014). However, the adoption of effective school-based tiered models of prevention requires universal screening for social, emotional, and behavioral health as a first step. As a preventative tool, screening data collected from students and teachers in the school setting permits school-based mental health professionals to: (a. better identify relevant universal programs and practices to combat problems at the universal or school/classroom-wide level [$>20\%$ experiencing a common risk factor]; (b. identify youth who may benefit from targeted supports due to experiencing elevated risk factors relative to their peers; (c. gain insight into the nature of universal and targeted problems to guide more in-depth assessment; (d. prioritize tasks necessary to support a population of students as well as individual students with greater needs

and compare screening scores across time to understand if school-wide efforts are reducing risk (Thompson et al., 2017).

In the past, some of the barriers cited by school personnel in opposition to school-wide screening practices for student social, emotional, and behavioral health included cost, time, lack of trained personnel, and lack of effective treatment options (Bruhn et al., 2014; Dowdy, Ritchey, & Kamphaus, 2010). Most notable among these barriers is the cost (Splett, Fowler, Weist, McDaniel, & Dvorsky, 2013). For example, the Behavioral and Emotional Screening System (BESS; Kamphaus & Reynolds, 2015), the Social Skills Improvement System (Elliott & Gresham, 2007), and the Systematic Screening for Behavioral Disorders (SSBD; Walker, Severson, & Feil, 2014) all cost about one dollar per student plus start-up costs (e.g., manuals, online scoring access, etc.). In addition, each of these systems does not effectively address a primary concern raised by school personnel in that too many children are identified by these screeners as requiring services (Volpe, Briesch, & Chafouleas, 2010). Because many screening systems rely on large national samples to establish norms and cutoff scores that are used to identify students who are at risk—we suggest these commercially available and nationally normed tools may overburden school personnel in high-risk geographical areas where resources are often limited. Although reliance upon national norms to establish statistical cutoff values is typical when developing screening tools, these practices do fail to take into account local culture and norms (Levitt, Saka, Romanelli, & Hoagwood 2007) as well as available resources to offer students effective supports.

Early Identification System

The barriers mentioned here led to the development of the Early Identification System (EIS). The EIS was developed by the *County Schools Mental Health Coalition* (the Coalition), a cooperative of six public districts, private schools, and university researchers, who worked to create the EIS to inform the deployment of a tiered prevention model. The Coalition developed the EIS to screen approximately 25,000 K-12 students across 55 school sites in a Midwestern county. After reviewing commercially available screening systems, the Coalition decided to create a tool that would avoid initial start-up and recurring costs of screening three times per year. Also, because the EIS was developed to be administered and scored online instantaneously, it takes high school teachers approximately 1 h to screen 150 students. Lastly, the EIS is also locally normed such that each student's risk score is derived in relation to her or his peers that they attend school alongside. The EIS provides a dashboard that permits school personnel to observe students who are in the top tier of risk (5%), students who are at moderate levels of risk (20%),

as well as patterns of risk by grade level and at the universal school level (Reinke, Sims, Cohen, & Herman, 2018a). The EIS includes both teacher and student report (EIS-TR, Grades K–12; EIS-SR, Grades 3–12) versions. These data can be used to prioritize needs and plan for universal prevention and/or targeted intervention strategies—which school personnel can select from using an EIS manual and accompanying Web site that lists specific evidence-based programs and practices previously shown to mitigate concerns identified by the screener. And lastly, because the EIS system is now the cornerstone of the Center for Rural Schools Mental Health (<https://www.ruralsmh.com/>) reporting on its psychometric qualities is important.

The EIS was initially developed with the guidance of the developmental cascades theory (Cicchetti & Tucker, 1994; Huang, Reinke, Thompson, Herman, & County Schools Mental Health Coalition, 2019; Patterson, Reid, & Dishion, 1992). According to the theory, the development of serious social, emotional, and behavioral health problems follows a predictive pattern—each event cascading into the next. Proposed by Patterson (1982; Patterson, DeBaryshe, & Ramsey, 1989), the model suggests children exposed to early coercive parenting at home start school with maladaptive schemas and scripts patterned upon escalating cycles of negative child–parent interactions (Patterson, 1982). These early patterns—dysregulation, aggression, attentional deficits—place youth at risk of failing to negotiate school-based tasks such as developing positive peer and teacher relations and performing well academically. Repeated failure and negative feedback at school tip youth into a cascade of negative downstream experiences intensify with key transitions (e.g., starting middle school). As such, consistent monitoring of these risks (externalizing, internalizing, attention and academics, peer relations, relational aggression, emotional dysregulation, and school disengagement) can help identify struggling students and select effective interventions to reduce the likelihood of poor outcomes.

Purpose of Present Study

The purpose of the present study is to confirm the factor structure of the EIS-SR in a large sample of public high school students (grades 9–12). For more contexts on the selection and development of the items and corresponding domains assessed by the EIS system, please see previously published factor analysis research reports by Huang et al. (2019); Reinke et al. (2018b) and Thompson et al. (2017). Initial investigation of the EIS factor structure using data from 1590 elementary school students (Huang et al., 2019) suggested the items fit well within the theorized factor structure with acceptable levels of scale reliability. Although a unified view of validity (Messick, 1995) considers other aspects of the instrument (e.g., the response process, the

interpretation, and consequences of the subscales) and is considered a continuing process, we focus here on a more basic level with three goals. First, using a split sample confirmatory factor analysis approach, would the EIS-SR replicate earlier observations of the factor structure in a pure high school sample? Next, would the EIS-SR have any measurement invariance by grade level and gender within the sample? Lastly, would EIS-SR scores predict later student-level discipline reports, attendance, bullying, and suspension data? Based upon earlier analyses with separate samples, we predicated the EIS-SR to factor similarly as observed, to be invariant with regard to grade and gender, and to be predictive of student outcomes.

Method

Participants

Participants (49.8% female) came from eight public high schools in a Midwestern state ($n = 5262$) with students in grades 9 ($n = 1580$; 30%), 10 ($n = 1382$, 26%), 11 ($n = 1258$, 24%), and 12 ($n = 1042$, 20%). Seventy-three percent of the students identified as White, 13% as Black, 5% as Asian, 5% as Latinx, and 5% as some other race or two or more races. Thirty-one percent of the students were eligible for free or reduced price meals (FRM; a commonly used proxy for socioeconomic status), and 12% had an identified disability.

Procedure

The EIS-SR was administered in the fall (October) of the school year and schools allowed parents to opt their child out of completing the EIS-SR. Students in grades 9–12 in participating schools completed the EIS-SR using a Web-based online format. A teacher (or a school mental health practitioner) read aloud a script about the purpose of the EIS-SR and how the information would be used. At the end of the survey, the school professional debriefed the students and provided them with contact information if they wanted to discuss the survey further with an adult. The students completed the measure in approximately 10 min.

Measures

The *Early Identification System-Student Report* (EIS-SR) was created through a process of item development and expert review. Authors first defined the purpose and scope of the scales based on an exhaustive literature review, generated items related to each factor, and placed each item within the corresponding factor. We generated a pool of 42 questions, and the item pool was reviewed by personnel in the schools as well as measurement experts. Response options

were Likert-type scales (0 = *Never*, 1 = *Sometimes*, 2 = *Often*, 3 = *Always*).

Seven factors were hypothesized according to Patterson's developmental cascades model: Externalizing Behavior, Internalizing Behavior, Peer Relationship Problems, School Disengagement, Emotional Dysregulation, Attention and Academic Issues, and Relational Aggression. Two items were deemed as stand-alone risk indicators (i.e., "Other kids make fun of me at school" and "I am bullied by others"), and thus, these items were excluded from the factor analyses. After interviews with students, one item was found to have questionable social validity as revealed by interviews with students ("I have friends to eat lunch with at school") and five other items (e.g., "I can solve real life problems") were excluded based on prior factor analytic work (Huang et al., 2019). In total, 34 items were hypothesized to form seven risk subscales.

Demographic Information

Schools provided demographic information for students, including gender, free and reduced (FRL) status, disability status, and race.

School Disciplinary Data

Schools provided office disciplinary referrals (ODRs) and the number of in-school (ISS) and out-of-school suspensions (OSS) received each month. For the purposes of this study, the total number of ODRs, ISS, and OSS was calculated based on those received across the months of January through the end of the school year. In the analytic sample, 22% received an ODR, 6% received an ISS, and 2% received an OSS in the spring.

Bullying Victimization

Students in the spring were asked "Other kids make fun of me at school" and "I am bullied by others" ($r = .39$). As bullying requires repetitive contacts (Solberg & Olweus, 2003), students who answered "Often" or "Always" were considered victims (2.5%).

Attendance Data

Schools provided the percent of days in attendance for each month. For spring attendance data, we averaged the monthly attendance data from January to May. Students who had attendance lower than 3 SD below the mean were excluded ($n = 15$).

Analytic Strategy

Data management and predictive validity analyses were conducted using R 3.5.3 (R Core Team, 2019). Factor analyses were performed using Mplus 6.1 (Muthen & Muthen, 2011). Prior to analysis, validity checks showed all responders suggested they had answered all of the items truthfully ($n = 4951$; 94%), a check shown to improve the quality of surveys (Cornell, Klein, Konold, & Huang, 2012; Furlong, Fullchange, & Dowdy, 2016; Jia, Konold, Cornell, & Huang, 2018). Analyses were conducted over three phases. First, we confirmed the factor structure of the EIS-SR (valid $n = 4951$). Second, we investigated the degree of measurement invariance by gender and grade level. Finally, we examined the predictive validity of the EIS-SR on outcomes.

Confirmatory Factor Analyses

Based on the researchers' knowledge of the items corresponding to the hypothesized factors and informed by a prior factor analytic study with a smaller sample Huang et al., 2019, we conducted a confirmatory factor analysis (CFA) using two randomly split samples of valid participants ($n_1 = 2475$, $n_2 = 2476$). Based on strong conceptual knowledge and prior research, researchers are suggested to perform a more stringent CFA rather than an exploratory factor analysis (Fabrigar & Wegener, 2012). To explore that factor structure within a CFA framework, we tested a unidimensional factor model and a seven-correlated-factor model. The one-factor model, if supported, would suggest that an overall measure of risk using all the items would be tenable. However, more realistically, we would expect that a seven (correlated)-factor model would fit better as we had hypothesized. We performed the CFA using the first random split-half sample (exploratory or training sample; $n = 2475$) and inspected resulting fit indices, factor loadings, and factor correlations. Items with low factor loadings (e.g., < 0.45) would be subject to removal (Comrey & Lee, 2013). Using the seven-correlated-factor model, if a factor exhibited extremely high correlations (e.g., > 0.80) with another factor, they would be subject to further inspection as well (e.g., factors may be combined). Using the exploratory sample and as a specification check, we consulted modification indices and items were evaluated based on their alignment with their theoretical postulates.

Given that a χ^2 test as a measure of model fit leads to the over rejection of reasonably specified models as a result of a large sample size (Anderson & Gerbing, 1988), several model fit indices were consulted. The fit indices included the root mean square error of approximation (RMSEA), the Tucker–Lewis index (TLI), and the comparative fit index (CFI). RMSEA values < 0.08 were considered reasonable

(Kline, 2011) and values for the CFI and TLI > 0.90 suggested acceptable model fit (Hu & Bentler, 1995).

Polychoric correlation matrices were used with weighted least squares with mean and variance correction (WLSMV) estimation were used to due to the categorical nature of the responses (Finney & DiStefano, 2006). To account for the clustered nature of the data, we used the type = complex option in Mplus which adjusted standard errors and model fit indices to account for the nesting of observations within schools (Stapleton, 2006).¹ To aid in model generalizability and avoid issues of overfitting, the best fitting model would then be replicated using another CFA but with the hold out sample. Scale score reliability was then estimated using categorical omega using the MBESS package (Kelley & Pornprasertmanit, 2016). Given the several limitations of Cronbach's alpha (e.g., biased, see Dunn, Baguley, & Brunson, 2014), we used omega which is interpreted in the same manner and "outperforms alpha and is clearly the preferred choice" (p. 7) for reporting scale reliability. To examine concurrent validity, we correlated subscale scores of the EIS-SR with each other. We considered correlations of approximately 0.50 to be large (Cohen, 1988; Kilgus, Chafouleas, & Riley-Tillman, 2013). Although we expected risk factors to exhibit some comorbidity with each other, we do not expect overly high correlations (i.e., $r > .80$) between the any two subscales.

Measurement Invariance

To investigate whether the instrument was measuring the same construct equally well among groups, measurement invariance was conducted (Dimitrov, 2010). We tested for measurement invariance using multi-group confirmatory factor analysis (MG-CFA) based on gender (male and female) and grade level (9, 10, 11, 12). The low numbers based on different racial classifications and the categorical nature of the data precluded the use of invariance testing based on race/ethnicity. Unlike gender or grade level which had over a thousand respondents in each subgroup, certain race/ethnicities (e.g., Asian, Latinx) had much than 300 total responses raising the likelihood that certain EIS questions did not have particular response endorsements (e.g., often or always) compared to White students, inhibiting the comparison of those questions.

Measurement invariance was conducted using three successively more restrictive models where configural, threshold, and loading invariance (Svetina, Rutkowski, &

Rutkowski, 2019; Wu & Estabrook, 2016) were investigated. Configural invariance (CI) tests whether the factors are measured by the same indicators across the specified groups (Bowen & Masa, 2015). Threshold invariance imposes an additional constraint where item thresholds, due to the ordinal nature of the data, are held equal among groups. Finally, loading invariance constrains the factor loadings to be equal among groups. Wu and Estabrook's (2016) invariance testing differs slightly from other methods (Bowen & Masa, 2015; Dimitrov, 2010) but is more optimal given the ordinal nature of the data (Svetina et al., 2019).

Several measures were inspected to determine whether invariance, or measurement equivalence was tenable using the succeeding more restrictive models (i.e., threshold vs. configural models, loading vs. threshold models) as the Satorra–Bentler (SB) χ^2 difference test is sensitive to sample size as well (Cheung & Rensvold, 2002). The Δ CFI is a commonly used invariance test where a decrease of CFI (i.e., a worsening of model fit) in the more restrictive model by more than 0.01 suggests noninvariance (Cheung & Rensvold, 2002). A Δ RMSEA of < 0.015 can be used as an indication of measurement invariance as well (Chen, 2007). Improvement in model fit may also provide evidence supporting measurement invariance (Dimitrov, 2010).

Predictive Validity

We used the fall EIS-SR subscales to predict whether a student would be bullied or receive an ODR, ISS, or OSS in the spring of the school year. EIS-SR subscales are scored such that higher scales represent greater risk (or more negative outcomes) and items were reverse scored as appropriate. Certain outcomes (e.g., ODRs) may signal a risk of future academic and behavioral problems (McIntosh, Flannery, Sugai, Braun, & Cochrane, 2008), and catching these risks early using subscales of EIS-SR would be helpful. We hypothesized that certain subscales (i.e., externalizing behaviors, emotional dysregulation) would predict disciplinary sanctioning. Attention issues and school disengagement would also likely predict future class attendance. In addition, we also used the EIS-SR subscales to predict class attendance in the spring of the school year. For bullying and disciplinary infractions, logistic regression models were used and for the continuous class attendance, a linear regression model was used. Using the entire sample, in all the regression models, student grade level, race/ethnicity, gender, disability status, and eligibility for free or reduced price meals were entered simultaneously in the models. All EIS-S subscales were standardized ($M = 0$, $SD = 1$). For the logistic regression analysis, the results are shown using odds ratios (ORs) values over 1 indicate higher odds of receiving the outcome (e.g., suspension or being bullied)

¹ Cluster robust standard errors are effective when the number of clusters is above 25 (Huang et al., 2019), using the current data, ICCs of the items were very low (2/3s of items having ICCs < 0.01 and 1 item approaching 0.05). Given the low ICCs as well, the clustering was judged as not problematic (Julian, 2001).

Table 1 Model fit statistics for one factor, multi-factor, and split sample

Model	χ^2	<i>df</i>	RMSEA	CFI	TLI
Exploratory sample (<i>n</i> = 2475)					
One-factor model	5094	527	0.059	0.860	0.851
Seven-correlated-factor model	1824	506	0.032	0.960	0.955
Hold-out sample (<i>n</i> = 2476)					
Seven-correlated-factor model	2070	506	0.035	0.950	0.945

DF degrees of freedom, *RMSEA* root mean square error of approximation, *CFI* comparative fit index, *TLI* tucker lewis index

and ORs lower than 1 indicate lower odds. To account for nesting, school fixed effects together with cluster robust standard errors were used (Huang, 2016).

Results

Factor Structure and Internal Consistency

Using the first randomly split-half sample (exploratory *n* = 2475), a one-factor model was tested using CFA (see Table 1). Fit indices indicated that the model did not fit well; RMSEA = 0.06, CFI = 0.76, TLI = 0.85. Next, a seven-correlated factor model was investigated and showed reasonable model fit; RMSEA = 0.03, CFI = 0.96, TLI = 0.96. To improve model generalizability, the seven-correlated factor CFA was also conducted using the second, confirmatory hold-out sample (*n* = 2476). Model fit using the hold-out sample was reasonable as well; RMSEA = 0.035, CFI = 0.95, TLI = 0.95. Standardized factor loadings using both samples are shown in Table 2. Factor loadings were strong with the weakest being 0.60. All items loaded on the hypothesized factors as theorized. Scale score reliability was assessed using omega using both samples. The subscale with the lowest reliability was for School Disengagement ($\omega = 0.71$), and the subscale with the highest reliability was for Internalizing Behavior ($\omega = 0.92$). Inspecting correlations between factors (see Table 3) shows a range of correlations from a low of 0.29 (for Relational Aggression with School Disengagement) to a high of 0.71 (for Externalizing Behavior and Attention and Academic Issues) and 0.74 (for Relational Aggression with Externalizing Behaviors). All of the risk subscales were positively correlated with each other indicating some level of comorbidity among the risk factors. None of the factors exhibited extremely high correlations (e.g., $r > .80$), suggesting that the factors were related but distinct from each other.

Measurement Invariance

Based on the procedures outlined for testing measurement invariance, the results (see Table 3) indicated that strong invariance was supported for both gender and grade level. As more restrictive models were tested by adding and testing equality of loadings, model fit did not deteriorate and all Δ RMSEA, Δ CFI, and Δ TLI statistics supported measurement invariance. In other words, regardless of gender or grade level, the EIS-SR measured the factors equally well.

Predictive Validity

To assess how well fall EIS-SR subscales would predict spring outcomes, we used all the EIS-SR subscales as predictors of disciplinary outcomes (i.e., receipt of ODR, ISS, OSS), bullying victimization, and attendance, while controlling for student demographic variables (Table 4).

Disciplinary Infractions

A consistent finding (see Table 5) was that students with higher levels of Externalizing Behavior (ORs = 1.21–1.27, $ps < .001$) and Emotional Dysregulation (ORs = 1.09–1.14, $ps < .05$) were more likely to receive a disciplinary infraction. In addition, students with higher Peer Relationship Problems (ORs = 0.92–0.96, $ps < .05$) and Internalizing Behaviors (ORs = 0.93–0.94, $ps < .001$) were less likely to receive a disciplinary infraction.

Bullying Victimization

The two subscales that were predictive of students being bullied in the spring were Internalizing Problems and Relational Aggression. Students with higher levels of Internalizing Problems (OR = 1.06, $p < .001$) and Relational Aggression (OR = 1.38, $p < .01$) were more likely to be victimized. All other EIS-SR subscales were not predictive of bullying victimization (Table 6).

Attendance

Students who experienced higher levels of Externalizing Behaviors ($\beta = -0.09$, $p < .01$), Attention Issues ($\beta = -0.12$, $p < .001$), and School Disengagement ($\beta = -0.07$, $p < .01$) were more likely to miss classes in the spring. Students with Peer Relationship Problems though were more likely to attend class ($\beta = 0.04$, $p < .05$) as well

Table 2 Standardized factor loadings using the exploratory ($n=2475$) and confirmatory, hold-out sample ($n=2476$)

Item	Question	PR	EXT	INT	ATT	ER	RA	SE
4	I am a good friend ^a	0.63 (0.70)						
5	I cooperate with others ^a	0.89 (0.85)						
6	I work well with my classmates	0.86 (0.88)						
18	I get in trouble at school		0.75 (0.76)					
19	I am sent out of class for bad behavior		0.81 (0.74)					
20	I disrupt class		0.74 (0.74)					
21	I get into fights with others		0.78 (0.78)					
22	My friends get in trouble at school		0.62 (0.62)					
31	I listen to my teachers ^a		0.76 (0.81)					
32	I blame others for my mistakes		0.73 (0.70)					
7	I have a hard time asking for help			0.65 (0.60)				
9	I like myself			0.73 (0.76)				
12	In the past month I felt sad			0.85 (0.84)				
13	In the past month I felt fearful			0.76 (0.75)				
14	In the past month I felt lonely			0.87 (0.85)				
15	In the past month I felt worried			0.79 (0.75)				
16	In the past month I felt like I did not matter			0.91 (0.91)				
17	In the past month I felt hopeless			0.90 (0.90)				
34	I feel left out by others			0.81 (0.79)				
23	I have trouble sitting still at school				0.71 (0.66)			
24	I have trouble finishing my work				0.79 (0.78)			
25	I have trouble paying attention				0.81 (0.84)			
29	I try hard to get good grades on my work ^a				0.62 (0.66)			
40	I complete my school work on time ^a				0.67 (0.62)			
26	I get mad easily					0.88 (0.89)		
27	I have a hard time controlling my temper					0.90 (0.84)		
33	I get crabby and irritated easily					0.82 (0.81)		
39	I need help with my emotions					0.87 (0.79)		
10	I am mean to others						0.96 (0.84)	
35	I talk about people behind their back						0.74 (0.69)	
36	I make fun of others						0.85 (0.85)	
28	I look forward to learning new things at school ^a							0.69 (0.67)
37	I enjoy coming to school ^a							0.78 (0.76)
42	There is an adult I can talk to at school if I need help							0.66 (0.65)
	Reliabilities (Omega)	0.77 (0.76)	0.76 (0.77)	0.92 (0.92)	0.86 (0.85)	0.85 (0.83)	0.74 (0.74)	0.71 (0.71)

Numbers within parenthesis represent statistics using the confirmatory, hold-out sample

PR peer relationships, EX externalizing behavior, IN internalizing behavior, AT attention issues, ED emotional dysregulation, RA relational aggression, SD school disengagement

^aReverse coded

as students with higher levels of Relational Aggression ($\beta=0.06$, $p < .05$). Both Internalizing Problems and Emotional Dysregulation measured in the fall were not predictive of attendance in the spring.

Discussion

Universal screening of social, emotional, and behavioral health is the key to identifying student needs and providing necessary support and prevention efforts within schools (Briesch, Chafouleas, & Chaffee, 2018; Dowdy et al., 2010). Unfortunately, many schools do not have the needed infrastructure, time, and school personnel support required to

Table 3 Factor correlations using exploratory sample (lower diagonal; $n = 2475$) and confirmatory sample (upper diagonal; $n = 2476$)

	PR	IN	EX	AT	RA	ED	SD
Peer relationship problems (PR)		0.49	0.57	0.53	0.60	0.54	0.51
Internalizing behavior (IN)	0.48		0.33	0.62	0.41	0.73	0.44
Externalizing behavior (EX)	0.56	0.37		0.71	0.74	0.55	0.42
Attention and academic issues (AT)	0.60	0.57	0.71		0.51	0.58	0.62
Relational aggression (RA)	0.52	0.37	0.64	0.49		0.55	0.33
Emotional dysregulation (ED)	0.56	0.69	0.48	0.55	0.50		0.39
School disengagement (SD)	0.55	0.43	0.47	0.58	0.29	0.38	

All correlations are statistically significant ($ps < .001$)

Table 4 Fit indices for invariance testing based on gender and grade level ($n = 4951$)

	χ^2	df	RMSEA	CFI	TLI
Gender (males = 2440; females = 2511)					
A. Configural invariance	3491	1012	0.031	0.962	0.958
B. Threshold invariance	3516	1046	0.031	0.962	0.959
C. Loading/threshold invariance	3514	1073	0.030	0.963	0.961
Grade level (grade 9 = 1508; grade 10 = 1291; grade 11 = 1176; grade 12 = 976)					
A. Configural invariance	4992	2024	0.034	0.958	0.953
B. Threshold invariance	5066	2126	0.033	0.958	0.956
C. Loading/threshold invariance	5077	2207	0.032	0.959	0.958

Satorra–Bentler $\Delta\chi^2$ testing Models B versus C were statistically significant ($ps < .001$). Based on Δ RMSEA, Δ CFI, Δ TLI, measurement invariance was supported

effectively screen their students. Coupled with the ever-expanding prevalence of mental health conditions in youth, this is extremely problematic. Thus, to address this problem, the EIS-SR was created as a brief and free assessment system to screen all students within a school. The EIS-SR was developed as a means of collecting accurate data to drive a tiered response model to support student social, emotional, and behavioral health. The EIS-SR was created based on extant literature of risk factors associated with social, emotional, and behavioral concerns among children and youth. The purpose of this study was to evaluate the factor structure, measurement invariance, and predictive validity of the EIS-SR. As anticipated, seven subscales of risk factors were revealed to have adequate factor loadings. This indicates that all items accurately reflect our hypothesized subscales—peer relation problems, internalizing behaviors, externalizing behaviors, attention and academic issues, relational aggression, emotional dysregulation, and school disengagement. All subscales also met reliability standards required for a screener (Cortina, 1993; Nunnally, 1978). In terms of internal consistency, all the risk subscales were found to be positively correlated with one another. However, no correlations were found to be extremely high, thus indicating that subscales were related, but still distinct from one another.

Beyond factor structure and internal consistency, the current study also evaluated invariance of subscales across

gender and grade level. The results of measurement invariance analyses found the same construct to be measured in a similar manner across both student gender and grade level. This is an important finding, as it is critical to ensure that the same underlying construct includes the same theoretical structure for each group while operating in a similar fashion. With invariance established, this allows the interpretations of the assessment results and subscale scores to be generalized to both female and male students and permits comparisons to be made between 9th, 10th, 11th, and 12th grade students. This is an incredibly helpful finding for schools, as it shows that the EIS-SR can be used to identify concerns and determine relevant supports for students in the same manner regardless of gender and grade level.

In addition, the current study assessed the predictive validity of the EIS-SR on student outcomes (i.e., disciplinary infractions, bullying victimization, and attendance). Demonstrating predictive validity of a measure on key external outcomes is an important step to establishing the meaning of a measure (Lord, Novick, & Birnbaum, 1968). For example, it makes sense that fewer ODRs in the fall would lead to fewer in and out of school suspensions—more time in class with fewer disruptions provides more opportunities for teachers to teach and students to learn.

Table 5 Logistic regression results (in odds ratios) predicting receipt of an office disciplinary referral (ODR), in-school suspension (ISS), and out-of-school suspension (OSS). 95% confidence intervals in parentheses

	ODR (<i>n</i> =4951)	ISS (<i>n</i> =4492)	OSS (<i>n</i> =4852)	Bullied (<i>n</i> =4951)
EIS subscales (standardized)				
Peer relations	0.95** (0.93, 0.98)	0.96* (0.92, 0.99)	0.92* (0.85, 1.00)	1.12 (0.98, 1.28)
Internalizing problems	0.94*** (0.93, 0.95)	0.94*** (0.92, 0.96)	0.93*** (0.90, 0.96)	1.06*** (1.02, 1.09)
Externalizing behaviors	1.22*** (1.17, 1.26)	1.27*** (1.19, 1.37)	1.21*** (1.13, 1.30)	1.00 (0.89, 1.13)
Attention issues	1.09*** (1.07, 1.12)	1.12** (1.04, 1.21)	1.07 (1.00, 1.15)	0.99 (0.91, 1.08)
Relational aggression	0.95 (0.88, 1.02)	0.87* (0.78, 0.98)	0.80*** (0.71, 0.91)	1.38** (1.11, 1.71)
Emotional dysregulation	1.09* (1.01, 1.16)	1.11* (1.02, 1.20)	1.14** (1.05, 1.25)	1.01 (0.91, 1.12)
School disengagement	1.09*** (1.06, 1.12)	1.05 (0.95, 1.15)	1.01 (0.95, 1.08)	1.03 (0.93, 1.14)
Student demographics				
Male	1.62*** (1.51, 1.73)	1.44** (1.11, 1.85)	1.62*** (1.22, 2.15)	1.30* (1.04, 1.63)
With a disability	0.94 (0.85, 1.04)	1.32** (1.10, 1.59)	1.52 (0.97, 2.38)	1.80*** (1.28, 2.54)
Eligible for FRM	2.03*** (1.71, 2.41)	3.58*** (2.80, 4.58)	3.52*** (1.87, 6.63)	1.19 (0.80, 1.77)
Race: Other ^a	1.43* (1.03, 1.98)	1.26 (0.91, 1.73)	1.80** (1.20, 2.72)	1.12 (0.74, 1.70)
Race: Asian ^a	0.57*** (0.51, 0.64)	0.23* (0.05, 0.94)	0.49 (0.14, 1.73)	1.04 (0.51, 2.11)
Race: Black ^a	2.40*** (2.17, 2.65)	2.39*** (1.67, 3.43)	2.50* (1.11, 5.64)	1.25 (0.76, 2.04)
Race: Hispanic ^a	1.16 (0.75, 1.79)	0.81 (0.33, 1.96)	1.77 (0.57, 5.43)	0.71 (0.18, 2.80)
Grade: 10 ^b	0.98 (0.70, 1.37)	1.11 (0.83, 1.50)	0.99 (0.62, 1.57)	1.18 (0.78, 1.80)
Grade: 11 ^b	0.92 (0.71, 1.19)	0.81 (0.49, 1.34)	0.66 (0.43, 1.01)	1.31 (0.78, 2.21)
Grade: 12 ^b	0.87 (0.59, 1.26)	0.81 (0.63, 1.05)	0.55** (0.35, 0.86)	0.43* (0.20, 0.93)

All outcomes and EIS scores standardized ($M=0$, $SD=1$). Cluster robust standard errors used. Models include school fixed effects

FRM free or reduced price meals

* $p < .05$; ** $p < .01$; *** $p < .001$

^aWhite is the reference group

^bGrade 9 is the reference group. Schools that had no recorded infraction (e.g., ISS or OSS) were excluded from the analysis

Implications

The needs of students with ever-increasing social, emotional, and behavioral health concerns continue to be unmet by many schools. This may be especially problematic in high school, as students with social, emotional, and behavioral

concerns have difficulties balancing academic demands, peer relationships, and increased independence, while also facing greater opportunities to engage in risky behaviors (Margherio, Evans, & Owens, 2019). Without appropriate services, adolescents are at increased risk for negative outcomes in high school (e.g., truancy, failing classes, and suspension/

Table 6 Linear regression results (cluster robust standard errors in parenthesis) predicting spring attendance ($n=4787$)

	Attendance
EIS subscales	
Peer relations	0.04* (0.01)
Internalizing problems	0.03 (0.03)
Externalizing behaviors	−0.09** (0.02)
Attention issues	−0.12*** (0.02)
Relational aggression	0.06* (0.02)
Emotional dysregulation	−0.04 (0.03)
School disengagement	−0.07** (0.02)
Student demographics	
Male	0.13** (0.03)
With a disability	−0.17* (0.05)
Eligible for FRM	−0.44*** (0.04)
Race: Other ^a	−0.09 (0.05)
Race: Asian ^a	0.26*** (0.01)
Race: Black ^a	0.02 (0.04)
Race: Hispanic ^a	−0.08 (0.08)
Grade: 10 ^b	−0.09 (0.05)
Grade: 11 ^b	−0.13 (0.06)
Grade: 12 ^b	−0.22** (0.06)
R^2	0.13

Attendance and EIS scores standardized ($M=0$, $SD=1$). Models include school fixed effects

FRM free or reduced price meals

* $p < .05$; ** $p < .01$; *** $p < .001$

^aWhite is the reference group

^bGrade 9 is the reference group. Outliers ($n=15$) were excluded and 3% of students had missing attendance records and were omitted

expulsion; Mitchell, Kern, & Conroy, 2019). Further, when compared to peers, high school students with significant depression are more than twice as likely to drop out of school (Dupéré et al., 2018). Social, emotional, and behavioral issues left untreated in high school can also persist and require more intensive services in adulthood (Heflinger, Shaw, Higa-McMillan, Lunn, & Brannan, 2015; Torio, Encinosa, Berdahl, McCormick, & Simpson, 2015). Universal screening is the first step to supporting these students. With effective universal screening in place, schools can make data-informed decisions and address student needs through intervention and prevention practices. Providing support for validation is the key to determining the utility of the EIS-SR in high school settings.

The EIS-SR is currently utilized within schools to identify, prevent, and address students' social, emotional, and behavioral issues—not only at the individual student level, but at the grade and school levels. EIS-SR data is locally normed within school buildings and allows individual students with the greatest level of need be identified without

straining school personnel or resources (Volpe et al., 2010). Based on results within a school building, individual student data can then be utilized within the context of problem-solving team to identify evidence-based interventions that target identified areas of concern (Reinke et al., 2018a, b). The grade-level results can be utilized in a similar fashion, as schools may determine preventive interventions to address needs present across an entire classroom. Further, data from the EIS-SR can be aggregated at the school level to determine if risk areas are present across many students. Based on a public health framework, the EIS-SR recommends that universal preventative interventions are implemented in schools when 20% of students or more indicate risk in a given area, as this approach is more productive than supporting each individual student separately (see Reinke et al., 2018a; Thompson et al., 2017).

Limitations and Future Directions

The current study is not without limitations. Although the current sample included over 5200 students across 8 high public high schools, all students were in the same geographic region of the country. Thus, it cannot be assumed that results are generalizable to students from other areas of the country. Additionally, over 70% of the sample was comprised of students who identified as White. Future studies should expand the use of the EIS-SR to more geographically and racially diverse contexts. The current study also did not evaluate concurrent validity of the EIS-SR in comparison with similar existing measures with robust psychometric properties (e.g., the SDQ or BASC-3). Future research correlating EIS-SR subscales with similar measures within a high school setting would further support the idea that the measure is accurately assessing what is expected. Lastly, the current study did not focus on creating cutoff scores for risk. Recently, researchers have noted the importance of systematically studying the implications of using cutoff scores within universal screening approaches (Margherio et al., 2019). Currently, the EIS-SR allows school personnel to identify individual students at the highest levels of risk based on school-normed data. However, future research identifying cutoff values may be beneficial to facilitating the classification of youth risk based on varying levels (e.g., low, moderate, high) that can be uniformly applied regardless of school-normed data. The results of the current study offer support for the EIS-SR as a free, brief, and effective tool to universally assess the social, emotional, and behavioral functioning of high school students. In particular, the present study established the factor structure, measurement invariance, and predictive validity of the EIS-SR. Compared to other comparable measures that have high costs, require trained school personnel, and/or are time-consuming, the EIS-SR offers an alternative that has potential for widespread use,

treatment utility, and impact. Further, the measure can be easily accessed, administered, and interpreted, thus allowing results to inform evidence-based intervention and prevention supports aimed to reduce the social, emotional, and behavioral concerns and disorders of students (see Reinke et al., 2018a, b; Thompson et al., 2017).

Funding The research reported here was supported by the Boone County Children's Services Fund and the US Department of Education, Institute of Education Sciences, #R305H170023 and #R305C190014. The opinions expressed are those of the authors and not the funders.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval The research performed here was approved and overseen by the Institutional Review Boards of the University of Missouri.

Human and Animal Rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and later amendments.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

- Anderson, M., & Cardoza, K. (2016). Mental health in schools: A hidden crisis affecting millions of students. *NPR Education*.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411. <https://doi.org/10.1037/0033-2909.103.3.411>.
- Arora, P. G., Collins, T. A., Dart, E. H., Hernández, S., Fetterman, H., & Doll, B. (2019). Multi-tiered systems of support for school-based mental health: A systematic review of depression interventions. *School Mental Health*, 11(2), 240–264.
- Barnett, S., Eber, L., & Weist, M. (2013). *Advancing education effectiveness: Interconnecting school mental health and school-wide positive behavior support*. Retrieved from <https://www.pbis.org/common/cms/files/Current%20Topics/Final-Monograph.pdf>.
- Battaglia, M., Garon-Carrier, G., Côté, S. M., Dionne, G., Touchette, E., Vitaro, F., et al. (2017). Early childhood trajectories of separation anxiety: Bearing on mental health, academic achievement, and physical health from mid-childhood to preadolescence. *Depression and Anxiety*, 34(10), 918–927.
- Boesen, M. L., & Lykke, K. (2012). Screening of vision and hearing in primary school children. *Journal of Family Medicine and Primary Care*, 1, 114–117.
- Bowen, N. K., & Masa, R. D. (2015). Conducting measurement invariance tests with ordinal data: A guide for social work researchers. *Journal of the Society for Social Work and Research*, 6(2), 229–249.
- Bradshaw, C. P., Reinke, W. M., Brown, L. D., Bevans, K. B., & Leaf, P. J. (2008). Implementation of school-wide positive behavioral interventions and supports (PBIS) in elementary schools: Observations from a randomized trial. *Education and Treatment of Children*, 31(1), 1–26.
- Briesch, A. M., Chafouleas, S. M., & Chaffee, R. K. (2018). Analysis of state-level guidance regarding school-based, universal screening for social, emotional, and behavioral risk. *School Mental Health*, 10(2), 147–162.
- Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education and Treatment of Children*, 37(4), 611–634.
- Chedekel, L. (2017). *Pediatric psychiatric cases continue climbing, swamping ERs*. Retrieved from <http://c-hit.org/2017/02/21/pediatricpsychiatric-cases-continue-climbing-swamping-ers/>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- Cicchetti, D., & Tucker, D. (1994). Development and self-regulatory structures of the mind. *Development and Psychopathology*, 6(4), 533–549.
- Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). New York, NY: Academic Press.
- Committee on Environmental Health [CEH]. (2005). Lead exposure in children: Prevention, detection, and management. *Pediatrics*, 116, 1036–1046.
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology press.
- Cornell, D., Klein, J., Konold, T., & Huang, F. (2012). Effects of validity screening items on adolescent survey data. *Psychological Assessment*, 24(1), 21–35. <https://doi.org/10.1037/a0024824>.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98–104. <https://doi.org/10.1037/0021-9010.78.1.98>.
- Curtin, S. C., & Heron, M. (2019). *Death rates due to suicide and homicide among person ages 10–24: United States, 2000–2017*. (NCHS Data Brief No. 352). Hyattsville, MD: National Center for Health Statistics.
- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43(2), 121–149. <https://doi.org/10.1177/0748175610373459>.
- Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health*, 2, 166–176.
- Ducharme, J. (2018). *A disturbing trend on the rise*. Retrieved from <http://time.com/5304227/suicide-on-the-rise/>.
- Dunn, T. J., Baguley, T., & Brunsdon, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105, 399–412. <https://doi.org/10.1111/bjop.12046>.
- Dupéré, V., Dion, E., Nault-Brière, F., Archambault, I., Leventhal, T., & Lesage, A. (2018). Revisiting the link between depression symptoms and high school dropout: Timing of exposure matters. *Journal of Adolescent Health*, 62, 205–211.
- Elliott, S. N., & Gresham, F. M. (2007). *Social skills improvement system: Classwide intervention program*. Bloomington, MN: Pearson Assessments.
- Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. New York, NY: Oxford University Press.
- Finney, S., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. Hancock & R. Mueller (Eds.), *Structural equation modeling: A second course*. Greenwich, CT: Information Age Publishing.

- Furlong, M. J., Fullchange, A., & Dowdy, E. (2016). *Effects of mischievous responding on universal mental health screening: I love rum raisin ice cream, really I do!* Retrieved from <http://psycnet.apa.org/psycinfo/2016-35905-001/>.
- Gersten, R., Clarke, B., Jordan, N. C., Newman-Gonchar, R., Haymond, K., & Wilkins, C. (2012). Universal screening in mathematics for the primary grades: Beginnings of a research base. *Exceptional Children, 78*, 423–445.
- Heflinger, C. A., Shaw, V., Higa-McMillan, C., Lunn, L., & Brannan, A. M. (2015). Patterns of child mental health service delivery in a public system: Rural children and the role of rural residence. *The Journal of Behavioral Health Services & Research, 42*, 292–309.
- Holen, S., Waaktaar, T., & Sagatun, Å. (2018). A chance lost in the prevention of school dropout? Teacher–student relationships mediate the effect of mental health problems on noncompletion of upper-secondary school. *Scandinavian Journal of Educational Research, 62*(5), 737–753.
- Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage Publications, Inc.
- Huang, F. L., Reinke, W. M., Thompson, A., Herman, K. C., & County Schools Mental Health Coalition. (2019). An investigation of the psychometric properties of the early identification system—student report. *Journal of Psychoeducational Assessment*. <https://doi.org/10.1177/0734282918758791>.
- January, S. A., Ardoin, S. P., Christ, T. J., Eckert, T. L., & White, M. J. (2016). Evaluating the interpretations and use of curriculum-based measurement in reading and word lists for universal screening in first and second grade. *School Psychology Review, 45*, 310–326.
- Jia, Y., Konold, T. R., Cornell, D., & Huang, F. (2018). The impact of validity screening on associations between self-reports of bullying victimization and student outcomes. *Educational and Psychological Measurement, 78*, 80–102. <https://doi.org/10.1177/0013164416671767>.
- Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling, 8*, 325–352.
- Kadhim, M., Lucak, T., Schexnayder, S., King, A., Terhoeve, C., Song, B., et al. (2020). Current status of scoliosis school screening: Targeted screening of underserved populations may be the solution. *Public Health, 178*, 72–77.
- 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.
- Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods, 21*(1), 69–92. <https://doi.org/10.1037/a0040086>.
- Kilgus, S. P., Chafouleas, S. M., & Riley-Tillman, T. C. (2013). Development and initial validation of the Social and Academic Behavior Risk Screener for elementary grades. *School Psychology, 28*, 210–226. <https://doi.org/10.1037/spq0000024>.
- Kline, R. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45*(2), 163–191.
- Lipkin, P. H., & Macias, M. M. (2020). Promoting optimal development: Identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics, 145*, 1–19.
- Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores*. Berlin: Addison-Wesley.
- Margherio, S. M., Evans, S. W., & Owens, J. S. (2019). Universal screening in middle and high schools: Who falls through the cracks? *School Psychology, 34*, 591.
- Maynard, B. R., Heyne, D., Brendel, K. E., Bulanda, J. J., Thompson, A. M., & Pigott, T. D. (2018). Treatment for school refusal among children and adolescents: A systematic review and meta-analysis. *Research on Social Work Practice, 28*(1), 56–67.
- McIntosh, K., Brigid Flannery, K., Sugai, G., Braun, D. H., & Cochrane, K. L. (2008). Relationships between academics and problem behavior in the transition from middle school to high school. *Journal of Positive Behavior Interventions, 10*, 243–255.
- McIntosh, K., Ty, S. V., & Miller, L. D. (2014). Effects of school-wide positive behavioral interventions and supports on internalizing problems: Current evidence and future directions. *Journal of Positive Behavior Interventions, 16*(4), 209–218.
- Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., et al. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the national comorbidity survey replication-adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980–989.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist, 50*, 741–749.
- Mitchell, B., Kern, L., & Conroy, M. (2019). Supporting students with emotional or behavioral disorders: State of the field. *Behavioral Disorders, 44*, 70–84.
- Muthen, L., & Muthen, B. (2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthen & Muthen.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Patterson, G. R. (1982). *Coercive family process* (Vol. 3). Castalia Publishing Company.
- Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychological Association, 44*(2), 329.
- Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys* (Vol. 4). Eugene: Castalia Pub Co.
- Power, T. J., Eiraldi, R. B., Clarke, A. T., Mazzuca, L. B., & Krain, A. L. (2005). Improving mental health service utilization for children and adolescents. *School Psychology, 20*, 187–205.
- R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from <https://www.R-project.org/>.
- Reinke, W. M., Sims, W., Cohen, D., & Herman, K. C. (2018a). Problem solving within an RTI framework: Roles and functions of effective teams. In Paige C. Pullen & Michael J. Kennedy (Eds.), *Handbook of response to intervention and multi-tiered systems of support* (pp. 80–103). New York: Routledge.
- Reinke, W. M., Thompson, A., Herman, K. C., Holmes, S., Owens, S., Cohen, D., et al. (2018b). The County Schools Mental Health Coalition: A model for community level impact. *School Mental Health, 10*, 173–180.
- Romer, D., & McIntosh, M. (2005). The roles and perspectives of school mental health professionals in promoting adolescent mental health. In D. L. Evans, E. B. Foa, R. E. Gur, H. Hendin, C. P. O'Brien, M. E. P. Seligman, & B. T. Walsh (Eds.), *Treating and preventing adolescent mental health disorders: What we know and what we don't know* (pp. 598–615). New York, NY: Oxford University Press.
- Saddler, B., & Asaro-Saddler, K. (2013). Response to intervention in writing: A suggested framework for screening, intervention, and progress monitoring. *Reading & Writing Quarterly, 29*, 20–43.
- Siceloff, E. R., Bradley, W. J., & Flory, K. (2017). Universal behavioral/emotional health screening in schools: Overview and feasibility. *Report on Emotional & Behavioral Disorders in Youth, 17*(2), 32.

- Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*(3), 611–625.
- Smith, J. C. (2019). The road to burnout: Dispositional and situational factors affecting teacher self-efficacy, job-related stress, and occupational burnout. *2019 NCUR*.
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior, 29*(3), 239–268. <https://doi.org/10.1002/ab.10047>.
- Splett, J. W., Fowler, J., Weist, M. D., McDaniel, H., & Dvorsky, M. (2013). The critical role of school psychology in the school mental health movement. *Psychology in the Schools, 50*(3), 245–258.
- Stapleton, L. (2006). Using multilevel structural equation modeling techniques with complex sample data. In G. Hancock & R. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 345–383). Greenwich, CT: Information Age Publishing.
- Stempel, H., Cox-Martin, M., Bronsert, M., Dickinson, L. M., & Allison, M. A. (2017). Chronic school absenteeism and the role of adverse childhood experiences. *Academic Pediatrics, 17*(8), 837–843.
- Svetina, D., Rutkowski, L., & Rutkowski, D. (2019). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using M plus and the lavaan/semTools Packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*, 111–130.
- Thompson, A. M., Reinke, W. M., Holmes, S., Danforth, L., Herman, K. C., & the County School Mental Health Coalition. (2017). The County School Mental Health Coalition: A model for a systematic approach to supporting youth. *Children & Schools, 39*, 209–218.
- Torio, C. M., Encinosa, W., Berdahl, T., McCormick, M. C., & Simpson, L. A. (2015). Annual report on health care for children and youth in the United States: National estimates of cost, utilization and expenditures for children with mental health conditions. *Academic Pediatrics, 15*, 19–35.
- Volpe, R. J., Briesch, A. M., & Chafouleas, S. M. (2010). Linking screening for emotional and behavioral problems to problem-solving efforts: An adaptive model of behavioral assessment. *Assessment for Effective Intervention, 35*, 240–244. <https://doi.org/10.1177/1534508410377194>.
- Walker, H. M., Severson, H. H., & Feil, G. E. (2014). *Systematic screening for behavior disorders (SSBD): Administrator's guide*. Eugene: Pacific Northwest Publishing.
- Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika, 81*(4), 1014–1045.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.