Hammill Institute on Disabilities

Journal of Emotional and Behavioral Disorders
2019, Vol. 27(2) 86–100

DOI: 10.1177/1063426617744746
jebd.sagepub.com
SAGE

Predictive Validity of *Student Risk Screening Scale for Internalizing and Externalizing* Scores in Secondary Schools

Kathleen Lynne Lane, PhD, BCBA-D[1], Wendy Peia Oakes, PhD[2], Emily D. Cantwell, M.Ed[1], David J. Royer, PhD[1], Melinda M. Leko, PhD[1], Christopher Schatschneider, PhD[3], and Holly Mariah Menzies, PhD[4]

Abstract

In this article, we examined predictive validity of *Student Risk Screening Scale for Internalizing and Externalizing* (SRSS-IE) scores for use at the middle (N = 2,313 from four middle schools) and high (N = 2,727 from two high schools) school level. Results indicated middle and high school students with high levels of risk according to fall SRSS-IE scores (particularly those with externalizing behaviors) were likely to have lower grade point averages, fail more courses, have more nurse visits, and spend more time in in-school suspensions compared with students at low risk for externalizing and internalizing behaviors. Education implications, limitations, and future directions are presented.

Keywords

systematic screening, tiered system of supports, at risk, middle school, high school

Middle and high school educational leaders across the country are recognizing the importance of meeting students' behavioral and social needs in addition to their academic needs (Watson, 2015; Yudin, 2014). This attention to behavioral and social supports is particularly encouraging given so many children and youth struggle with externalizing and internalizing behavior (Forness, Freeman, Paparella, Kauffman, & Walker, 2012). Externalizing behaviors often include aggressive, noncompliant, and hostile tendencies which are quick to capture teachers' attention as these behaviors frequently disrupt learning environments. In contrast, internalizing behaviors are often more covert in nature, often including shy, anxious, and social withdrawal tendencies. Although students with internalizing behaviors may not be disruptive to learning environments, these behaviors are no less serious and can be challenging for students and society as a whole as they negatively affect relationships with others and academic outcomes (Bradshaw, Buckley, & Ialongo, 2008; Green et al., 2017; Lane & Walker, 2015). Furthermore, both of these major disorders are far more common than one might expect.

Recent point prevalence estimates offered by Forness and colleagues (2012) clearly established that many school-age youth experience externalizing and/or internalizing behavior patterns. They reported 20% of school-age youth have mild-to-severe emotional and behavioral disorders (EBD), with 80% of these challenges manifesting before they leave high school (Forness et al., 2012). The magnitude of EBD is troublesome given the negative associated outcomes for this group of students: lack of school connectedness, school failure, in-grade retention, school dropout, strained interpersonal relationships, under- and unemployment, and increased need for mental health supports (Maggin, Wehby, Farmer, & Brooks, 2016; Shochet, Dadds, Ham, & Montague, 2006; Siperstein, Wiley, & Forness, 2011; Wagner, 1995). Decades of research has clarified students do not "outgrow" externalizing or internalizing behaviors. Left unchecked, these challenges persist over time well into and beyond middle and high school. Furthermore, the costs of the associated deleterious outcomes are high for these individuals, their families, and society as a whole (Farmer et al., 2015; Walker, Forness, & Lane, 2014).

When one considers the vast number of adolescents struggling with internalizing and externalizing behaviors, the tendency for these challenges to persist over time in the absence of evidence-based interventions, the peer rejection

[1]University of Kansas, Lawrence, USA
[2]Arizona State University, Tempe, USA
[3]Florida State University, Tallahassee, USA
[4]California State University, Los Angeles, USA

Corresponding Author:
Kathleen Lynne Lane, Department of Special Education, College of Education, University of Kansas, 1122 West Campus Road, JRP Room 542, Lawrence, KS 66045, USA.
Email: Kathleen.Lane@ku.edu

associated with and the negative outcomes characteristic of these behavior challenges, it is critical to build psychometrically sound, feasible tools for conducting systematic screenings. It is particularly important to construct these tools for use in secondary schools as students tend to shift away from externalizing attempts to seek peer and social acceptance and instead develop responses to social rejection or isolation that may result in demonstrations of violence, self-injurious behaviors including substance abuse, suicidal ideation, and other internalizing concerns such as depression and anxiety. This developmentally typical shift in the importance of peer acceptance may place students at risk for externalizing and internalizing behaviors in a vulnerable position given their behaviors often negatively affect relationships with peers as well as adults (Farmer et al., 2015; Lane, Oakes, Carter, Lambert, & Jenkins, 2013; Maggin et al., 2016; Walker, Forness, & Lane, 2014). With educational funding always a consideration for many school districts, it is critical for the research community to explore psychometrically sound, free-access, and low-cost approaches for detecting secondary students with challenging behaviors—the purpose of this study.

Fortunately, many astute district- and school-site leaders are acknowledging the value of a systems perspective to prevent and respond to students' academic, behavior, and social needs in a holistic manner (McIntosh & Goodman, 2015). Several schools and districts are focusing on how to design, implement, and evaluate comprehensive, integrated, three-tiered (Ci3T) models of prevention to effectively and efficiently meet students' multiple needs (Lane, Oakes, & Menzies, 2014). Ci3T models provide a cascade of empirically validated supports for primary (Tier 1) prevention efforts for all, secondary (Tier 2) prevention efforts for some, and tertiary (Tier 3) prevention efforts for a few. Central to this—and any—tiered system of supports is accurate detection of which students require more than primary (Tier 1) prevention efforts have to offer (Oakes, Lane, Cox, & Messenger, 2014).

Through the years, several screening tools initially developed and validated for use in elementary settings have been expanded and refined for use in middle and high schools (e.g., Walker, Severson, & Feil, 2014). The expanded range of screening tools for use with secondary-age students in detecting internalizing and externalizing behaviors is encouraging, as is the availability of free-access tools given the uncertainty of the future of educational funding. One such free-access, teacher-completed tool is the *Student Risk Screening Scale for Internalizing and Externalizing behaviors* (SRSS-IE; Drummond, 1994; Lane & Menzies, 2009; see ci3t.org for a copy of the SRSS-IE). The SRSS-IE is an adapted version of the SRSS (Drummond, 1994) designed to detect elementary-age students with antisocial behaviors. In the last 10 years, the original seven-item tool has been expanded to detect internalizing issues, adding five new items yielding the SRSS-IE12: (1) steal; (2) lie, cheat, sneak; (3) behavior problem; (4) peer rejection; (5) low academic achievement; (6) negative attitude; (7) aggressive behavior; (8) emotionally flat; (9) shy, withdrawn; (10) sad, depressed; (11) anxious; and (l2) lonely. Teachers independently screen students on their designated class roster by completing each item using the same 4-point Likert-type scale developed by Drummond (1994): *never* = 0, *occasionally* = 1, *sometimes* = 2, and *frequently* = 3.

Although the SRSS-IE was initially developed for use at the elementary level (e.g., Lane, Menzies, et al., 2012; Lane, Oakes, et al., 2012; Lane et al., 2015), recent inquiry in middle and high schools offers preliminary evidence of the utility of the SRSS-IE with secondary-age students (Lane, Oakes, Carter, et al., 2013). For example, Lane et al. (2017) conducted a psychometric study in which they reported initial evidence of the SRSS-IE scores in middle (n = 9) and high schools (n = 3) from three states. In this first examination of the SRSS-IE in high schools, results of an exploratory factor analysis of SRSS-IE scores suggested five of the seven proposed items designed to measure internalizing behavior patterns be retained for use with secondary-age students, yielding the SRSS-IE12. Retained items at the secondary level were the same five items retained for use in elementary schools: emotionally flat; shy, withdrawn; sad, depressed; anxious; and lonely. Yet, results indicated the item *peer rejection* (one of the originally developed items by Drummond, 1994), loaded on the internalizing construct, representing a shift in the role of peer rejection.

Whereas peer rejection tended to be more reflective of externalizing behavior patterns in the elementary years (Lane, Oakes, et al., 2012), peer rejection appeared to be more characteristic of internalizing issues in middle and high schools (Lane et al., 2017). This finding is consistent with work by Farmer and colleagues who have examined the complex nature of social relationships in children and youth. Namely, preschool and elementary students who experience peer rejection may use aggressive behaviors to improve their social status with peers. Farmer and colleagues further found that students who experienced peer rejection and demonstrated aggressive behaviors had higher social status than those with shy or nonaggressive behavior, even when they may not be liked by peers (Farmer, Farmer, Estell, & Hutchins, 2007). Teachers rated students with higher social status as having fewer internalizing concerns (Farmer, Hall, Leung, Estell, & Brooks, 2011). Conversely, Farmer et al. (2011) found more than half of students who experienced peer rejection were identified with low social status, putting them at risk for further social rejection and isolation. Furthermore, students who experience social rejection with low social prominence were more apt to experience victimization and other negative social interactions (Farmer et al., 2011) which show strong causal

relations with internalizing concerns (e.g., depression, anxiety, self-injurious behaviors; Moore et al., 2017).

This preliminary evidence proposed two factors: SRSS-E7 (as originally designed) and SRSS-I6 for use with secondary students, with *peer rejection* loading on both subscales (but only added once when computing the total score for the SRSS-IE12). This important finding regarding the dual role of *peer rejection* was consistent with the work of Farmer et al. (2011) who have continued to explore peer rejection. Their research suggested there are various types of peer rejection (e.g., victimization, levels of social prominence, isolation; Farmer et al., 2007; Farmer et al., 2011). Farmer and colleagues (2011) reported teachers are able to distinguish various facets (i.e., levels of social prominence) of peer rejection manifesting in late childhood and early adolescence. More specifically, in considering the relationship between externalizing behaviors, internalizing behaviors, and peer rejection, it is important to understand the distinction between being disliked by one's peers (peer rejection) and being popular (perceived popularity) as youth progress from childhood through adolescence (Cillessen & Rose, 2005; Rodkin, 2011). When students are young, peer rejection in the elementary years is typically associated with externalizing behaviors. Students who exhibit behaviors characteristic of externalizing behavior disorders are often viewed as unpopular and are definitely not perceived as being leaders (Dawes et al., 2017). However, as students progress through late childhood and into early adolescence there is a shift. Externalizing behaviors appear to be associated with still being disliked, but now popular (e.g., consider mean girls or athletes who bully; referred to as populistic—popular, but not well liked; de Bruyn & Cillessen, 2006). Yet, for students who exhibit behaviors characteristic of internalizing behavior disorders, they tend to withdrawal from social relationships as they transition into adolescences. These students may be viewed by their teacher as being rejected by their adolescent peers and are less likely to attribute them as having aggressive behaviors (Dawes et al., 2017; Rodkin, 2011).

After exploring the factor structure of the SRSS-IE in middle and high schools, Lane, Oakes, Cantwell, Schatschneider, et al. (2016) conducted another study yielding preliminary cut scores for the SRSS-IE12 to facilitate data-informed decision making in middle and high schools. Specifically, they reported findings of a convergent validity study examining the internalizing subscale (SRSS-I6) from the SRSS-IE12 with the internalizing subscale of the Teacher Report Form (TRF; Achenbach, 1991). This sample included 227 sixth- through 12th-grade students from nine schools across two states. Results of receiver operating characteristic (ROC) curves and logistic regressions yielded the following cut scores for the SRSS-I6 with secondary-age students: 0–3 low risk, 4–5 moderate risk, and 6–18 high risk for internalizing behavior patterns (Lane, Oakes, Cantwell, Schatschneider, et al., 2016).

Findings of these studies provided initial evidence of the score reliability and validity. However, inquiry has not been conducted to examine predictive validity. Predictive validity refers to the degree to which a score on a scale predicts scores on a given criterion measure. For example, to what extent do fall SRSS-E7 and SRSS-I6 scores predict important behavioral and academic outcomes for students?

The next step in this programmatic line of inquiry is to examine the predictive validity of SRSS-IE subscale scores with this secondary-age population. Previous inquiry at the middle and high school levels indicated SRSS-E7 scores predicted office discipline referrals (ODRs), suspensions, grade point averages (GPA), and even course failures. For example, at the middle school level, short-term predictive validity examined with 500 sixth- through eighth-grade students suggested low-, moderate-, and high-risk status was most differentiated by in-school suspensions and ODRs. Furthermore, students without (low) and with (moderate and high) risk were differentiated by GPA and number of course failures, with students with low-risk status having the more successful outcomes (Lane, Parks, Kalberg, & Carter, 2007). Findings were confirmed by a series of studies conducted by Lane, Bruhn, Eisner, and Kalberg (2010) offering additional evidence of long-term predictive validity, with fall SRSS scores predicting student outcomes up to 2 years later. Specifically, students with moderate- and high-risk scores had more unexcused absences, were more likely to be suspended, and ended the year with lower GPAs than students in the low-risk group.

Similarly, at the high school level, original SRSS scores predicted important short-term and long-term outcomes for students (Lane, Kalberg, Parks, & Carter, 2008; Lane, Oakes, Ennis, et al., 2013). In the first study of the SRSS with 674 ninth- through 12th-grade students, Lane et al. (2008) reported predictive validity over two academic years. Students with low-risk status could be differentiated from students in moderate- or high-risk categories on ODRs and GPA. Lane, Oakes, Ennis, et al. (2013) replicated this study with 1,854 high school students. Results established predictive validity of SRSS scores across two academic years, with spring scores differentiating students with low-, moderate-, and high-risk status on ODRs, GPA, and course failures.

At this time, psychometric studies of SRSS-IE scores have provided preliminary evidence of the SRSS-IE factor structure (Lane et al., 2017) and cut scores (Lane, Oakes, Cantwell, Schatschneider, et al., 2016) as applied with middle and high school students. We now seek to determine the extent to which SRSS-I6 scores also predict important outcomes for middle and high school students.

Purpose

The intent of this study was to provide initial evidence to support the utility of SRSS-IE scores for use in secondary schools, following the data analytic plan developed by Lane et al. (2007) to explore predictive validity of the original SRSS as applied to middle school students. Specifically, we present two studies in which validity of SRSS-IE scores were examined in middle (Study 1, $n = 4$) and high (Study 2, $n = 2$) school settings. In each study, we examined predictive validity of fall SRSS-IE scores by analyzing the extent to which middle and high school students with low, moderate, and high risk for externalizing (SRSS-E7 scores) and internalizing (SRSS-I6 scores) could be differentiated in terms of behavioral and academic characteristics according to extant schoolwide data. Specifically, these studies (a) replicate previous inquiry establishing predictive validity of SRSS-E7 scores and (b) explore the predictive validity of SRSS-I6 scores in secondary schools. At the middle school level, we examined GPA, course failures, nurse visits, in-school suspensions, and ODRs. At the high school level, we examined these same variables, with the exception of ODRs as these data were not available. Consistent with earlier inquiry, we hypothesized SRSS-E7 and SRSS-I6 scores would be more indicative of behavioral rather than academic outcomes as the former is more closely aligned with constructs assessed using the SRSS-IE behavior screening tool (Lane et al., 2007).

Study 1: Method

Participants and Setting

Participants were 2,313 middle school students (1,202 males) attending one of four middle schools in the Midwest who were rated by their advisory teachers ($n = 171$) on the SRSS-IE. Schools served students in sixth (34.52%), seventh (33.81%), and eighth (31.67%) grades. Students were predominantly White (75.08%, $n = 1,690$), with approximately 22.35% of students receiving special education services (see Table 1). Economic disadvantage rates varied across schools, ranging from 21.91% to 58.22% (see Table 2 for school characteristics).

Procedures

Each middle school established a Ci3T Leadership Team comprised of the principal, two general education teachers, a special education teacher, two to three other individuals (e.g., instructional coaches, counselors, and school psychologists), a parent, and student. Each team attended a year-long professional learning series offered by the district in partnership with two university partners to design, implement, and evaluate a Ci3T model of prevention. As part of their assessment plan, faculty who provided instruction to students completed the SRSS-IE 3 times over the course of the school year: 6 weeks after the school year began (fall), before winter break (winter), and 6 weeks prior to year-end (spring). Data were used by each school's Ci3T Leadership Team and Ci3T District Leadership Team members to (a) examine the overall level of behavioral risk evident in each school, (b) inform the use of teacher-delivered low-intensity supports to facilitate student engagement (e.g., instructional choice, increasing opportunities to respond), and (c) connect students to Tier 2 and 3 supports (e.g., Robertson & Lane, 2007). All schools were in their first year of implementing Ci3T. As part of their initial implementation efforts, they first focused on implementing primary (Tier 1) efforts. Specifically, school-site Ci3T Leadership Teams first worked with the Ci3T District Leadership Team to learn how to analyze data to examine overall levels of risk in the building and then inform teacher-delivered supports such as behavior-specific praise and increasing opportunities to respond. They were in the beginning stages of learning how to use screening data in conjunction with regular school practices to connect students with appropriate Tier 2 and Tier 3 supports.

Ci3T Leadership Teams introduced the SRSS-IE to their respective faculty in the spring of their Ci3T training year during a regularly scheduled faculty meeting to familiarize them with the purpose of screening and offer an opportunity to learn more about administration and interpretation. Middle school principals and district leaders determined that Advisory teachers would conduct screenings for their students. Advisory period was 20 to 25 min daily (depending on the schedule) and was the period selected for social skills instruction. School leaders felt this context allowed for teachers to get to know their students in a broader context than any individual content area and enabled more teachers to participate in the screening activities. Teachers completed all items on the SRSS-IE. However, each Ci3T Leadership Team understood the majority of their decision making would focus on the SRSS-E7 subscale score during the first year of implementation as the internalizing subscale was under development for middle and high school students at that time (see "Measures" section for recent psychometric properties of the internalizing subscale). Data presented in the current study are from the 2015–2016 academic year, the district's first year of implementing Ci3T across all secondary schools (four middle and two high schools).

The district established a secure method to enable teachers to complete the SRSS-IE independently, yielding one independent teacher rating for each student. Student names and identification numbers were prepopulated for each teacher's advisory class approximately 30 days before each screening window opened in fall, winter, and spring. Teachers were electronically permissioned to access the SRSS-IE data sheet for only their one assigned period. Ci3T Leadership Team members explained to teachers total scores (not item-level data) on the SRSS-E7 would be used

Table 1. Student Characteristics.

	Middle	High
Variable/level	*N* = 2,313	*N* = 2,727
Gender % (*n*)		
Male	53.40 (1,202)	52.89 (1,393)
Female	46.60 (1,049)	47.11 (1,241)
Grade % (*n*)		
Sixth	34.52 (777)	—
Seventh	33.81 (761)	—
Eighth	31.67 (713)	—
Ninth	—	29.23 (770)
10th	—	23.99 (632)
11th	—	25.06 (660)
12th	—	21.72 (572)
Ethnicity/race % (*n*)		
Hispanic	8.17 (184)	7.97 (210)
White	75.08 (1,690)	76.23 (2,008)
Black	6.75 (152)	6.45 (170)
Asian/Pacific Islander	3.91 (88)	2.92 (77)
Native American/Native Alaskan	4.31 (97)	4.02 (106)
Declined	0.49 (11)	0.38 (10)
Mixed races	9.46 (213)	9.98 (263)
Special education % (*n*)	22.35 (503)	19.89 (524)
Emotional disturbance	10.23 (35)	11.85 (39)
Intellectual disability	2.92 (10)	4.56 (15)
Speech language delays	8.77 (30)	1.52 (5)
Learning disabilities	50.29 (172)	50.46 (166)
Autism spectrum disorder	10.23 (35)	11.55 (38)
Other health impaired	14.62 (50)	15.50 (51)

Note. N represent all students enrolled over the course of the academic year. Data are reported for ethnicity (i.e., Hispanic) and race for students. Special education eligibility reported for categories with 10 or more students assigned.

Table 2. School Characteristics.

	School					
Variable	MS A	MS B	MS C	MS D	HS A	HS B
Grades served	6–8	6–8	6–8	6–8	9–12	9–12
Teachers completing screeners	36	43	46	46	84	77
Attendance rate[a]%	94	94.8	95.3	94.9	93.7	93.4
Graduation rate[a]%	—	—	—	—	92.6	91.4
FRPL[b]%	58.22	53.04	21.91	38.74	29.29	41.71
Title I eligible[b]	Yes	Yes	No	Yes	No	Yes

Note. Locale—City: Small. FRPL= free and reduced-price lunch eligible; MS = middle school; HS = high school.
[a]State school report card data 2014–2015. [b]National Center for Education Statistics, Common Core Data 2014–2015.

for decision making until secondary level cut scores were available for the internalizing subscale (SRSS-I6).

The Ci3T District Leadership Team shared de-identified, student-level data electronically with principal investigators. Data were coded and checked for accuracy prior to data analysis. In this paper, we report findings of fall 2015 screening data in predicting five year-end student outcomes: GPA, course failures, nurse visits, in-school suspensions, and ODRs.

Measures

SRSS-IE. The SRSS-IE is an adapted version of the SRSS, expanded to include additional items characteristic of

internalizing behaviors. The original seven items (SRSS-E7) developed by Drummond (1994) to detect antisocial behaviors were retained in exact form: steal; lie, cheat, sneak; behavior problems; peer rejection; low academic achievement; negative attitude; and aggressive behavior. Teachers independently rate each behavior on a 4-point Likert-type scale developed by Drummond (1994) of *never* = 0, *occasionally* = 1, *sometimes* = 2, and *frequently* = 3, yielding the following cut scores for the SRSS-E7: 0–3 low risk, 4–8 moderate risk, 9–21 high risk. Following a series of psychometric studies, five of the seven proposed internalizing items (emotionally flat; shy, withdrawn; sad, depressed; anxious; lonely) were added to the SRSS to form the SRSS-IE12 (Lane, Oakes, Carter, et al., 2013; Lane, Oakes, Ennis, et al., 2013). Convergent validity analyses of SRSS-IE scores with elementary and middle school students suggested SRSS-IE scores have convergent validity with Strengths and Difficulties Questionnaire (SDQ) scores (Lane, Oakes, Carter, et al., 2013). Furthermore, evidence from a psychometric study exploring the factor structure of SRSS-IE scores in nine middle schools and three high schools from three states with more than 10,000 students suggested a two-factor solution (Lane et al., 2017). The *peer rejection* item (initially developed by Drummond, 1994) loaded more prominently on the internalizing construct resulting in two factors: SRSS-E7 and SRSS-I6 for secondary students, with the peer rejection item loading on both factors (yet added once when computing the total score, SRSS-IE12). Results of the first convergent validity study examining the internalizing subscale (SRSS-I6) scores with the internalizing subscale scores of the TRF (Achenbach, 1991) suggested the following cut scores for the SRSS-I6 for sixth through 12th grades: 0–3 low risk, 4–6 moderate risk, and 7–18 high risk (Lane, Oakes, Cantwell, Schatschneider, et al., 2016). In this article, we used current cut scores for SRSS-E7 and SRSS-I6 subscale scores to examine predictive validity.

Extant schoolwide data. We predicted the following year-end outcomes: GPA, course failures, nurse visits, in-school suspensions, and ODRs. District leaders provided de-identified year-end data electronically to principal investigators. GPA referred to cumulative performance on a 4-point scale. Course failures referred to the total number of Ds or Fs earned during the same academic year. Nurse visits referred to the total number of visits a student made to the office for assistance of any kind (e.g., getting a bandage, nausea, somatic complaints). In-school suspensions referred to the total number of days a student was assigned in-school suspension due to serious or repeated infractions. We defined ODRs as the number of ODRs each student earned according to the school's information system. Each Ci3T Leadership Team defined a schoolwide reactive plan as part of their Ci3T model of prevention, listing behavioral examples that resulted in students earning an ODR. A series of logic checks were completed to ensure data provided by the district reflected accurate ranges.

Statistical Analysis

Students were grouped into low, moderate, and high levels of risk as described previously. Groups of students were then examined to investigate potential differences in GPA, course failures, nurse visits, in-school suspensions, and ODRs. To investigate potential differences in GPA by group, we fit a mixed-model ANOVA with group as a fixed effect and classroom teacher as a random effect. This model takes into account the nested nature of the data (students nested within teachers' classes) to determine the degree to which students scoring in the low-, moderate-, and high-risk categories according to fall SRSS-E7 and SRSS-I6 scores could be differentiated on GPA. Significant group effects were followed up with a set of pairwise comparisons. The Type I error-rate for these post hoc tests was controlled using a Bonferroni correction, with the alpha level for each group comparison set at .05/3 = .0167.

For the dependent variables measured as counts, we conducted a series of random-effects negative binomial regressions with an over dispersion parameter. These models also take into account the nested nature of the data to determine the extent to which students with low, moderate, and high levels of risk according to fall SRSS-E7 and SRSS-I6 scores could be differentiated on course failures, nurse visits, in-school suspensions, and ODRs collected over the course of the 2015–2016 academic year. We chose to fit negative binomial regression models for all variables except GPA given the distributions of these outcome variables more closely resemble a Poisson distribution which is commonly seen in count variables. The negative binomial regression that models over dispersion is particularly useful for samples in which there are many people with zeros (e.g., zero course failures, zero nurse visits, and zero ODRs). This model is most appropriate when dependent variables are distributed as count data and the standard deviation of the count variable exceeds the mean count (which was the case for these data). Analyses were completed using data provided; missing data were not imputed, but missingness was handled via full maximum likelihood estimation for the negative binomial regressions and mixed-model ANOVAs (Enders, 2010).

Finally, we computed effect sizes from observed means and standard deviations to determine the magnitude of differences between groups using the Hedges's *g* formula (using the pooled standard deviation in the denominator). Effect sizes were interpreted as follows: small- (0.20), medium- (0.50) and large-magnitude effects (0.80; Cohen, 1988).

Results

SRSS-E7 at the Middle School Level

Findings of a mixed-model ANOVA with group as the between-subjects fixed effect and teacher as the random effect indicated a group effect for GPA, $F(2, 1872) = 211.00$, $p < .0001$ (see Table 3). The low-risk externalizing group

Table 3. Middle School: Behavioral and Academic Characteristics of Risk Groups According to Fall SRSS-IE Subscale Scores.

	Risk				Effect size		
	Low	Moderate	High				
Subscale/variable	*M* (*SD*) *n*	*M* (*SD*) *n*	*M* (*SD*) *n*	Significance testing	L:M	L:H	M:H
Externalizing							
Grade point average	3.56 (0.47) 1,670	3.07 (0.58) 279	2.74 (0.61) 84	L > M > H	1.01	1.72	0.56
Course failures	0.38 (1.15) 1,830	1.37 (2.12) 328	2.78 (3.03) 93	L < M < H	0.74	1.84	0.60
Nurse visits	4.01 (16.20) 1,830	6.67 (8.65) 328	9.66 (11.65) 93	L < M, H M = H	0.17	0.35	0.32
Office discipline referrals	0.03 (0.24) 1,830	0.17 (0.63) 328	0.75 (2.13) 93	L < M < H	0.42	1.38	0.51
In-school suspensions	0.11 (0.89) 1,830	0.67 (2.74) 328	1.56 (3.22) 93	L < M < H	0.42	1.30	0.31
Internalizing							
Grade point average	3.51 (0.51) 1,642	3.33 (0.55) 167	3.16 (0.64) 224	L > M > H	0.35	0.66	0.28
Course failures	0.52 (1.42) 1,820	0.86 (1.85) 181	1.22 (2.06) 250	L < M, H M = H	0.23	0.46	0.18
Nurse visits	4.32 (16.39) 1,820	4.85 (6.92) 181	6.77 (9.56) 250	L < H L = M M = H	0.03	0.16	0.22
Office discipline referrals	0.06 (0.40) 1,820	0.17 (1.24) 181	0.19 (0.75) 250	*ns*	0.21	0.29	0.02
In-school suspensions	0.18 (1.10) 1,820	0.67 (3.59) 181	0.45 (1.47) 250	L < M, H M = H	0.33	0.24	−0.09

Note. SRSS-IE = Student Risk Screening Scale for Internalizing and Externalizing Behaviors; H = high risk; L = low risk; M = moderate risk; *ns* = post hoc comparisons suggest no statically significant differences. Confidence Intervals available from authors.

earned a statistically significantly higher GPA than moderate- ($t = 15.62$, $p < .0001$, Hedges's $g = 1.01$) and high-risk groups ($t = 15.16$, $p < .0001$, Hedges's $g = 1.72$). The moderate-risk group had a statistically significantly higher mean GPA score than the high-risk group ($t = 5.58$, $p < .0001$, Hedges's $g = 0.56$).

For number of course failures, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 76.58$, $p < .0001$. Post hoc comparisons revealed the low risk for externalizing group experienced significantly fewer course failures than moderate- ($t = 9.80$, $p < .0001$, Hedges's $g = .74$) and high-risk groups ($t = 9.34$, $p < .0001$, Hedges's $g = 1.84$). The moderate-risk group experienced fewer course failures than the high-risk group ($t = 3.05$, $p < .0023$, Hedges's $g = .60$).

For number of nurse visits, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 44.70$, $p < .0001$. Post hoc comparisons revealed the low risk for externalizing group experienced significantly fewer nurse visits than moderate- ($t = 7.54$, $p < .0001$, Hedges's $g = 0.17$) and high-risk groups ($t = 6.82$, $p < .0001$, Hedges's $g = 0.35$). The difference in nurse visits between moderate- and high-risk groups for externalizing behaviors was not statistically significant ($t = 2.24$, $p = .0250$, Hedges's $g = 0.32$).

For number of ODRs, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 38.86$, $p < .0001$. Post hoc comparisons indicated the low risk for externalizing group experienced significantly fewer ODRs than moderate- ($t = 5.85$, $p < .0001$, Hedges's $g = 0.42$) and high-risk groups ($t = 8.23$, $p < .0001$, Hedges's $g = 1.38$). The moderate-risk group experienced fewer ODRs than the high-risk group ($t = 3.84$, $p = .0001$, Hedges's $g = 0.51$).

For number of in-school suspensions, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 63.97$, $p < .0001$. Post hoc comparisons revealed the low risk for externalizing group experienced significantly fewer in-school suspensions than moderate- ($t = 8.47$, $p < .0001$, Hedges's $g = 0.42$) and high-risk groups ($t = 9.70$, $p < .0001$, Hedges's $g = 1.30$). Furthermore, students in the moderate-risk group earned fewer in-school suspensions than students in the high-risk groups ($t = 3.65$, $p = .0003$, Hedges's $g = 0.31$).

SRSS-I6 at the Middle School Level

Findings of a mixed-model ANOVA indicated a group effect for GPA, $F(2, 1872) = 43.59$, $p < .0001$. The low-risk internalizing group earned a statistically significantly higher GPA than moderate- ($t = 4.29$, $p < .0001$, Hedges's $g = 0.35$) and high-risk groups ($t = 8.86$, $p < .0001$, Hedges's $g = 0.66$). The moderate-risk group had a statistically significantly higher mean GPA than the high-risk group ($t = 2.92$, $p = .0036$, Hedges's $g = 0.28$).

For number of course failures, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 16.57$, $p < .0001$. Post hoc comparisons revealed the low risk for internalizing group experienced significantly fewer course failures than both moderate- ($t = 3.06$, $p = .0023$, Hedges's $g = 0.23$) and high-risk groups ($t = 5.36$, $p < .0001$, Hedges's $g = 0.46$). There was no statistically significant difference in course failures between moderate- and high-risk groups ($t = 1.23$, $p = .2198$, Hedges's $g = 0.18$).

For number of nurse visits, we fit a random-effects negative binomial regression model, demonstrating a significant overall omnibus test, $F(2, 2078) = 16.50$, $p < .0001$. Post hoc comparisons revealed a statistically significant difference between the low- and high-risk groups ($t = 5.64$, $p < .0001$, Hedges's $g = 0.16$), with the low-risk group having fewer nurse visits. There were no statistically significance differences in nurse visits between low- and moderate-risk groups ($t = 2.00$, $p = .0459$, Hedges's $g = 0.03$), nor were there statistically significant differences between moderate-risk and high-risk groups ($t = 2.37$, $p = .0181$, Hedges's $g = 0.22$) after taking into account the Bonferroni correction with a .0167 alpha criterion.

For number of ODRs, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 3.01$, $p = .0495$, barely meeting the .05 criterion. Post hoc comparisons revealed no statistically significant differences in ODRs between low-, moderate-, and high-risk groups.

For number of in-school suspensions, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2078) = 9.80$, $p < .0001$. Post hoc comparisons revealed the low risk for internalizing group experienced significantly fewer in-school suspensions than both moderate- ($t = 2.87$, $p < .0001$, Hedges's $g = 0.33$) and the high-risk groups ($t = 3.87$, $p = .0042$, Hedges's $g = 0.24$). There was no statistically significant difference in the number of days of in-school suspension between moderate- and high-risk groups ($t = 0.42$, $p = .6723$, Hedges's $g = -0.09$).

Study 2: Method

Participants and Setting

Participants were 2,727 students (1,393 males) attending one of two public high schools in the Midwest who were rated by their second-period teachers ($n = 161$) on the SRSS-IE. Schools served students in ninth (29.23%), 10th (23.99%), 11th (25.06%), and 12th (21.72%) grades. Students were predominantly White (76.23%, $n = 2,008$), with 19.89% of students receiving special education services (see Table 1). Economic disadvantage rates were 29.29% for high school A and 41.71% for high school B (see Table 2 for School Characteristics).

Procedures

Procedures for Study 2 were identical to Study 1, with the exception of the Ci3T Leadership Team composition, screening period, and variables predicted. At the high schools, teams also included a second administrator (e.g., vice principal or administrative intern), two parents, and two student members, given the size of the schools. Second-period teachers were selected by building principals and district leaders to complete the SRSS-IE because some students were not on campus during first period. Second-period course offerings included the full scope of classes (including core and elective courses) offered in traditional high schools. Screening teachers were representative of the full scope of courses offered at the high schools. ODR data were not predicted in this study as these data were not available.

Results

SRSS-E7 at the High School Level

Findings of a mixed-model ANOVA indicated a group effect for GPA, $F(2, 2479) = 172.62$, $p < .0001$ (see Table 4). The low-risk externalizing group earned a statistically significantly higher average GPA than moderate- ($t = 16.64$, $p < .0001$, Hedges's $g = 1.25$) and high-risk groups ($t = 9.96$, $p < .0001$, Hedges's $g = 1.40$). There was no statistically significant difference in average GPA scores between moderate- and high-risk groups ($t = 1.25$, $p = .2123$, Hedges's $g = 0.15$).

For number of course failures, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2479) = 46.72$, $p < .0001$. Post hoc comparisons revealed the low risk for externalizing group experienced significantly fewer course failures than both the moderate- ($t = 9.05$, $p < .0001$, Hedges's $g = 1.05$) and high-risk groups ($t = 4.32$, $p < .0001$, Hedges's $g = 0.92$). There was no statistically significant difference in the number of course failures between

Table 4. High School: Behavioral and Academic Characteristics of Risk Groups According to Fall SRSS-IE Subscale Scores.

	Risk						
	Low	Moderate	High	Significance testing	Effect size		
Subscale/variable	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)		L:M	L:H	M:H
Externalizing (*n*)	2,363	212	59				
Grade point average	3.07 (0.79)	2.08 (0.81)	1.96 (0.89)	L > M, H M = H	1.25	1.40	0.15
Course failures	1.16 (2.07)	3.45 (3.18)	3.08 (2.84)	L < M, H M = H	1.05	0.92	−0.12
Nurse visits	1.34 (3.19)	4.00 (5.62)	5.85 (7.66)	L < M, H M = H	0.77	1.34	0.30
In-school suspensions	0.07 (0.44)	0.67 (1.48)	1.03 (1.86)	L < M, H M = H	1.00	1.84	0.23
Internalizing (*n*)	2,379	123	132				
Grade point average	3.04 (0.82)	2.44 (0.83)	2.27 (0.98)	L > M, H M = H	0.73	0.93	0.19
Course failures	1.25 (2.17)	2.59 (2.66)	2.83 (3.21)	L < M, H M = H	0.61	0.71	0.08
Nurse visits	1.43 (3.33)	3.54 (6.05)	4.04 (5.80)	L < M, H M = H	0.60	0.75	0.08
In-school suspensions	0.11 (0.57)	0.41 (1.36)	0.42 (1.28)	L < M, H M = H	0.48	0.49	0.01

Note. SRSS-IE = Student Risk Screening Scale for Internalizing and Externalizing Behaviors; H = high risk; L = low risk; M = moderate risk Confidence Intervals available from authors.

moderate- and high-risk groups ($t = 0.47$, $p = .6403$, Hedges's $g = -0.12$).

For number of nurse visits, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2479) = 52.08$, $p < .0001$. Post hoc comparisons suggested the low-risk externalizing group accessed significantly fewer nurse visits than moderate- ($t = 8.59$, $p < .0001$, Hedges's $g = 0.77$) and high-risk groups ($t = 6.28$, $p < .0001$, Hedges's $g = 1.34$). There was no statistically significant difference in the average number of nurse visits between moderate- and high-risk groups ($t = 1.42$, $p = .1564$, Hedges's $g = 0.30$).

For number of in-school suspensions, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2479) = 62.88$, $p < .0001$. Post hoc comparisons indicated the low risk for externalizing group experienced significantly fewer in-school suspensions than both the moderate-risk group ($t = 9.98$, $p < .0001$, Hedges's $g = 1.00$) and the high-risk group ($t = 6.71$, $p < .0001$, Hedges's $g = 1.84$). There was no statistically significant difference in the number of days of in-school suspension between moderate- and high-risk groups ($t = 1.00$, $p = .3186$, Hedges's $g = 0.23$).

SRSS-I6 at the High School Level

Findings of a mixed-model ANOVA indicated a group effect for GPA, $F(2, 2479) = 68.43$, $p < .0001$. The low-risk internalizing group had a statistically significant higher GPA than moderate- ($t = 7.09$, $p < .0001$, Hedges's $g = 0.73$) and high-risk groups ($t = 9.89$, $p < .0001$, Hedges's $g = 0.93$). Difference in GPA between the moderate- and high-risk groups was not statistically different ($t = 2.03$, $p = .0429$, Hedges's $g = 0.19$).

For number of course failures, we fit a random-effects negative binomial regression model. The model demonstrated a significant overall omnibus test, $F(2, 2479) = 25.52$, $p < .0001$. Post hoc comparisons revealed the low risk for internalizing group experienced significantly fewer course failures than both moderate- ($t = 4.93$, $p < .0001$, Hedges's $g = 0.61$) and high-risk groups ($t = 5.57$, $p < .0001$, Hedges's $g = 0.71$). There was no statistically significant difference in the number of course failures between moderate- and high-risk groups ($t = .32$, $p = .7477$, Hedges's $g = 0.08$).

For number of nurse visits, we fit a random-effects negative binomial regression model which demonstrated a significant overall omnibus test, $F(2, 2479) = 30.79$, $p < .0001$. Post hoc comparisons revealed the low risk for internalizing group experienced significantly fewer nurse visits than moderate- ($t = 5.27$, $p < .0001$, Hedges's $g = 0.60$) and high-risk groups ($t = 6.34$, $p < .0001$, Hedges's $g = 0.75$). There was no statistically significant difference in the number of nurse visits between moderate- and high-risk groups ($t = 0.66$, $p = .5112$, Hedges's $g = 0.08$).

For number of in-school suspensions, we fit a random-effects negative binomial regression model. The model

demonstrated a significant overall omnibus test, $F(2, 2479) = 12.02$, $p < .0001$. Post hoc comparisons revealed the low risk for internalizing group experienced significantly fewer in-school suspensions than moderate- ($t = 3.73$, $p = .0002$, Hedges's $g = 0.48$) and high-risk groups ($t = 3.55$, $p = .0004$, Hedges's $g = 0.49$). There was no statistically significant difference in the number of days of in-school suspension between moderate- and high-risk groups ($t = 0.06$, $p = .9526$, Hedges's $g = 0.01$).

Discussion

Given the magnitude of students with internalizing and/or externalizing behavior patterns, the persistent nature of these challenges in the absence of evidence-based interventions implemented with high fidelity, the peer rejection experienced in secondary schools, and less than optimal postsecondary outcomes for these students, it is essential to focus on building effective, efficient, systematic screening tools for use in secondary schools (Farmer et al., 2015; Maggin et al., 2016; Walker, Forness, & Lane, 2014). The focus on psychometrically sound accessible tools becomes increasingly important when confronted with the need to efficiently use often limited educational funding. Again, we reiterate that the intent of this study was to examine the SRSS-IE, not to promote it as "better" than any of the existing commercially available tools. We examined the utility of SRSS-IE scores in predicting important educational outcomes for secondary-age students.

Specifically, the intent of the two studies reported in this article was to provide initial evidence to support the utility of SRSS-IE scores for use in secondary schools. To this end, we reported outcomes examining validity of SRSS-IE scores in middle (Study 1) and high (Study 2) school settings, with a particular emphasis on determining the degree to which middle and high school students with low, moderate, and high risk for externalizing (SRSS-E7 scores) and internalizing (SRSS-I6 scores) could be differentiated on measures of behavioral and academic outcomes. Thus, while secondary teachers have less contact time with students than do elementary teachers, findings show SRSS-IE scores predict important outcomes for students at the secondary level.

SRSS-E7: Predictive Validity of Externalizing Scores in Middle and High Schools

In predicting Year 1 outcomes at the middle school level, fall SRSS-E7 scores differentiated low-, moderate-, and high-risk groups on GPA, course failures, ODRs, and in-school suspensions. The low-risk group had statistically significantly higher year-end GPAs as well as fewer course failures, ODRs, and in-school suspensions compared with moderate- and high-risk groups. Furthermore, moderate- and high-risk group mean scores could also be differentiated, with students in the high-risk group having the most deleterious outcomes. Findings were similar to results of previous short-term (1 year) predictive validity studies also suggesting fall SRSS-E7 scores predicted GPA and course failures (Lane et al., 2010; Lane et al., 2007). However, this was the first study to examine the extent to which screening scores predicted nurse visits. In this study, we learned fall SRSS-E7 scores differentiated students in the low-risk group relative to students in the moderate- and high-risk groups for externalizing behaviors. Students in the low-risk group had fewer nurse visits than students in the moderate- or high-risk groups for externalizing behaviors. There were no statistically significant differences in nurse visits for students in the moderate- or high-risk, suggesting any heightened risk for externalizing behaviors put students at heightened risk for nurse visits. We note this information should be considered preliminary in nature until this work is replicated. Yet, number of nurse visits is an important outcome to explore as frequent nurse visits may indicate a range of needs and challenges such as medical assistance stemming from an altercation, social anxiety (e.g., not wanting to enter a class late, not having a lunch group to join, issues with teacher or peers, bullying), lack of school connectedness (feeling no one truly cares), or other unmet needs such as food, clothing, or attention (Johnson & Hutcherson, 2006; Vernberg, Nelson, Fonagy, & Twemlow, 2011). By providing students with interventions targeting the underlying needs of their nurse visits, school nurses can then attend to medical, case management, and other duties more efficiently and effectively. This is a significant need given the National Association of School Nurses recommends a nurse to student ratio of 1:750 in a healthy population and lower in populations with more complex health needs, but this is often not the case due to shortages of school nurses and/or school budget cuts (American Association of Colleges of Nursing, 2014; Holmes et al., 2016).

In examining effect sizes, results indicated high-magnitude effects when differentiating low- and high-risk groups on GPA (1.72), course failures (1.84), ODRs (1.38), and in-school suspensions (1.30). However, effect sizes were medium-to-large when differentiating nurse visits between low- and high-risk groups (.77). Collectively, results indicated SRSS-E7 scores continue to be an effective screening tool for predicting not only behavioral but also academic outcomes at the middle school level.

In terms of utility of SRSS-E7 scores at the high school level, findings reported here closely paralleled earlier inquiries reporting short-term (1 year) as well as long-term (2 years) predictive validity. Lane et al. (2008) found students with low-risk status could be differentiated from students in moderate- or high-risk groups on ODRs and GPA as applied with a sample of 674 ninth- through 12th-grade students. Lane, Oakes, Ennis, et al. (2013) replicated this study with 1,854 high school students. Findings established predictive

validity of SRSS-E7 scores across two academic years, with spring scores differentiating students with low-, moderate-, and high-risk on ODRs, GPA, and course failures. In the current sample (as was the case with Lane, Oakes, Ennis, et al., 2013), there is a distinction in risk between fall mean scores for high school students placing into moderate- and high-risk groups for externalizing behaviors as evidenced by statistical significance testing as well as clear high-magnitude differences between the low-risk group relative to moderate- or high-risk groups. Although not statistically significant in this sample, some effect sizes suggested small-to-medium distinctions between moderate- and high-risk groups at the high school. This suggests any sign of risk—moderate or high—should be attended to efficiently and carefully in secondary schools at the fall time point. Although these outcomes are consistent with fall predictive validity outcomes previously reported (Lane, Oakes, Ennis, et al., 2013), we note spring SRSS-E7—and for some variables, winter SRSS-E7 scores—could be differentiated between moderate- and high-risk status. We strongly suggest additional inquiry be conducted to replicate these predictive validity analyses prior to generalizing these results in other locales and with students from diverse socioeconomic backgrounds. It is critical researcher and practitioner communities not prematurely conclude risk is simply a dichotomous variable (low vs. any risk) when examining outcomes. Instead, we echo previous recommendations noting the lack of distinction between moderate- and high-risk scores. In other words, fall SRSS-E7 scores may suggest any level of risk at the onset of an academic year is cause for concern and may warrant additional consideration within (and outside of) a tiered system of supports (Lane, Oakes, Ennis, et al., 2013; Walker, Severson, & Feil, 2014). To the maximum extent possible, when examining fall screening scores, we encourage school leadership teams to determine how to best support students in the moderate- and high-risk groups at the first sign of concern given these scores predicted important academic and behavioral outcomes in middle and high school (McIntosh, Frank, & Spaulding, 2010). We offer the same recommendation when interpreting internalizing behavior patterns.

SRSS-I6: Predictive Validity of Internalizing Scores in Middle and High Schools

In this first predictive validity study of SRSS-I6 scores, middle school results suggested sixth- through eighth-grade students at low, moderate, and high risk for internalizing behaviors could also be differentiated on GPA. Students in the low-risk group for internalizing behavior could also be differentiated on course failures from students in the moderate- and high-risk groups. However, the moderate- and high-risk groups could not be differentiated. This finding supports the complex relation between internalizing behaviors and successful school experiences as students with internalizing behaviors may experience challenges with concentration, school engagement, and other self-determined behaviors needed for school success (Fröjd et al., 2008; Shochet et al., 2006). As has often been discussed, strong interpersonal skills are necessary to develop the requisite skills for successful negotiation of relationships with peers and adults—referred to as academic enablers (Malecki & Elliot, 2002; Rapport, Denney, Chung, & Hustace, 2001; Walker, Irvin, Noell, & Singer, 1992). Students at high risk for internalizing behaviors could also be differentiated on nurse visits from the low-risk group; however, the other contrasts were not statistically significant following the Bonferroni correction. We note the contrast between moderate- and high-risk groups came very close to meeting criterion. While the increasing number of nurse visits for students in the high risk category was expected, replication is particularly critical in this area to better understand possible manifestations of internalizing behavior patterns in secondary schools. As discussed, the frequent need for nurse visits is troublesome on multiple levels as it may reflect a range of concerns: experiencing discomfort, unmet physical and/or emotional needs, missing instruction, and potentially reinforcing escape-motivated behaviors (e.g., somatic complaints; Moore et al., 2017). In addition, at the middle school level, students at high-risk levels earned more days of in-school suspension relative to students in the low-risk group. This may be due to the complex nature and challenges associated with establishing social structures at the middle school level when students with internalizing behaviors may be most vulnerable to involvement in peer problems such as bullying resulting in social conflict and aggression (Farmer et al., 2011; Farmer et al., 2015; Rodkin, 2011). For example, when instances of bullying occur, all involved parties might be disciplined (possibly resulting in in-school suspensions)—particularly when people are provoked to the point of verbal and physical aggression. Although, we note effect sizes indicated the magnitude of the differences between low- and high-risk internalizing groups were of a lower magnitude than the differences between externalizing groups for in-school suspensions, as would be expected. Consistent with McIntosh, Campbell, Carter, and Zumbo's (2009) finding of no statistically significant relation between ODRs and ratings of internalizing behaviors at the elementary level, our findings from Study 1 suggested low-, moderate-, and high-risk groups for internalizing behaviors at the middle school level were not differentiated on ODRs. Yet, again we call for replication as the overall model for ODRs barely met the .05 criterion.

Interestingly, fall SRSS-I6 scores performed in the same manner as SRSS-E7 scores at the high school level. Fall internalizing scores could differentiate students with low-risk status from students in moderate- or high-risk

categories on course failures, nurse visits, and in-school suspensions. This was evidenced by statistical significance testing as well as clear medium- to large-magnitude effect sizes when comparing low-risk groups to moderate- or high-risk groups. In addition, there was no distinction between moderate- and high-risk group scores for internalizing behaviors for course failures, nurse visits, and in-school suspensions as evidenced by statistical tests and effect size calculations.

However, it could be predictive validity may be more precise when using winter and spring screening data, similar to findings with SRSS-E7 scores (Lane, Oakes, Ennis, et al., 2013). This may be more true for internalizing behaviors given their covert nature (Farmer et al., 2011; McIntosh et al., 2010; Walker, Forness, & Lane, 2014) and the relation between adolescents' school connectedness and increasing internalizing behaviors (Shochet et al., 2006). As mentioned, early inquiry at the high school level demonstrated SRSS-E7 scores from subsequent administrations demonstrated improved predictive validity compared with fall scores. Lane, Oakes, Ennis, and colleagues (2013) hypothesized predictive validity improved given teachers had more time to interact with students—particularly in year-long courses and less structured (e.g., study-hall type) periods—resulting in additional time to learn more about their academic, behavioral, and social performance patterns. For example, in the Lane, Oakes, Ennis et al. study of long-term predictive validity, spring SRSS-E7 scores at the end of the first year differentiated all three groups of students (low-, moderate-, and high-risk status) on GPA, course failures, and ODRs. Students with high-risk status failed more classes, earned lower year-end GPAs, and acquired more ODRs than students in the moderate-risk group who failed more classes, earned lower GPAs, and earned more ODRs than students in the low-risk group during the *following* academic year. This same pattern of responding may hold true for SRSS-I6 scores as well.

Educational Implications

Additional information on the utility of SRSS-E7 scores and the preliminary nature of the utility of SRSS-I6 scores suggested SRSS-IE scores are useful for distinguishing between students in the low-risk group from students in the moderate- and high-risk groups on most academic and behavioral variables examined in this study. It will be important for future inquiry to determine if additional time with students (e.g., winter and/or spring scores) will continue to improve predictive validity of SRSS-E7 and SRSS-I6 scores collected at later time points. It will be especially important to determine if year-end internalizing scores are more accurate in predicting student outcomes 1 year later. Given the negative outcomes for students with these difficult to detect behaviors, this is a key point for future inquiry.

When interpreting fall screening scores, we encourage school leadership teams and individual teachers to examine multiple sources of data and develop a plan for supporting students in moderate- and high-risk groups for externalizing and/or internalizing issues at the earliest possible juncture given the negative outcomes likely to occur. For secondary students in particular, school leadership teams might consider multiple access points for students to develop relationships with school personnel and become actively engaged in the school community. School connectedness serves as a protective factor for adolescents (Shochet et al., 2006). Support may also focus on the use of teacher-delivered strategies such as increased use of behavior-specific praise, incorporating instructional choices, and increasing student opportunities to respond (Jolivette, Stichter, & McCormick, 2002; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). Then, for students whose risk status does not improve by the next behavior screening, it might be prudent to move to evidence-based practices such as check in/checkout or other self-management strategies (see Carter, Lane, Crnobori, Bruhn, & Oakes, 2011; Lane, Menzies, Bruhn, & Crnobori, 2011, for examples). It will be important to provide high-quality professional learning to assist teacher acquisition of the knowledge and confidence to use these strategies as well as see their utility and feasibility within the context of their regular instructional routines. The goal is to empower all faculty and staff with the skillsets they need to differentiate instruction and incorporate positive behavioral interventions and supports (PBIS) as part of their normal instructional activities as opposed to seeing "behavior" as something to be managed separately from instruction (McIntosh & Goodman, 2015; Menzies, Lane, Oakes, & Ennis, 2017). Given the evidence from this study suggests any indication of risk predicts important academic and behavioral outcomes, we must move forward with a comprehensive, integrated approach to meeting students' multiple needs.

Limitations and Future Directions

Although we are pleased to share these preliminary predictive validity results with the researcher and practitioner communities, we encourage readers to interpret results with attention to the following limitations. First, replication is critical before generalizing these results to other middle and high school contexts (e.g., locales, regions, socioeconomic status). Although this study included data from four middle and two high schools, the sample is from one district in one locale with Ci3T prevention models in place. We invite other research teams to conduct similar studies in other locales, particularly those serving more ethnically and culturally diverse communities before drawing definitive conclusions.

Second, this study did not address issues of comorbidity. Although not a goal of the present study, we encourage future inquiry to explore the predictive validity of SRSS-IE scores taking into account issues of comorbidity. Namely, it would be important to know if students who experience low, moderate, or high risk on both externalizing (SRSS-E7) and internalizing (SRSS-I6) dimensions could also be differentiated on these and other student performance measures.

Third, as in previous psychometric studies at the middle and high school, participating schools were part of a researcher–practitioner partnership grant funded through the Institute for Education Sciences (IES), with a strong interest in installing and utilizing systematic screening within Ci3T models of prevention being implemented at their respective schools. Future research is needed to explore the degree to which these patterns of responding are similar when screenings are selected, implemented, and utilized in isolation from researcher–practitioner partnerships and with various other school structures in place (Lane, Oakes, Ennis, et al., 2013).

Summary

Despite limitations, results of the two studies presented extended the knowledge base with respect to predictive validity of SRSS-E7 (measuring externalizing behaviors) and SRSS-I6 (measuring internalizing behaviors) scores for predicting a range of academic and behavioral outcomes for middle and high school students. Results indicated students with high levels of risk (particularly those with externalizing behaviors), as measured by the SRSS-IE just 4 to 6 weeks after the school year began, were likely to have lower GPAs, fail (D or F grades) more courses, have more nurse visits, and spend more time in in-school suspensions compared with students at low risk for externalizing and internalizing behaviors. Although we are pleased to offer this first study of the predictive validity of SRSS-IE scores in middle and high schools with a large sample, we urge readers to use this information cautiously to avoid generalization errors until replication studies are completed. Nonetheless, this is an important first step to suggest one independent rating of student performance using the SRSS-IE is clearly able to distinguish between students with low- and high-risk status for externalizing and internalizing behaviors on a range of academic and behavioral outcomes in middle and high schools.

Authors' Note

Opinions expressed herein are those of the authors and do not necessarily reflect the position of the U.S. Department of Education, and such endorsements should not be inferred. David J. Royer is presently affiliated to University of Hawai'i at Mānoa.

Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by the Office of Special Education Programs U.S. Department of Education (H326S980003) and Institute of Education Sciences Partnership Grant (R305H150018).

ORCID iD

David J. Royer https://orcid.org/0000-0003-2882-1049

References

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YRS, and TRF profiles*. Burlington: Department of Psychiatry, University of Vermont.

American Association of Colleges of Nursing. (2014). *Nursing shortage*. Retrieved from http://www.aacn.nche.edu/media-relations/fact-sheets/nursing-shortage

Bradshaw, C. P., Buckley, J. A., & Ialongo, N. S. (2008). School-based service utilization among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly*, *23*, 169–186. doi:10.1037/1045-3830.23.2.169

Carter, E. W., Lane, K. L., Crnobori, M. E., Bruhn, A. L., & Oakes, W. P. (2011). Self-determination interventions for students with and at risk for emotional and behavioral disorders: Mapping the knowledge base. *Behavioral Disorders*, *36*, 100–116.

Cillessen, A. H. N., & Rose, A. J. (2005). Understanding popularity in the peer system. *Current Directions in Psychological Science*, *14*, 102–105. doi:10.1111/j.0963-7214.2005.00343.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Dawes, M., Chen, C.-C., Zumbrunn, S. K., Mehtaji, M., Farmer, T. W., & Hamm, J. V. (2017). Teacher attunement to peer-nominated aggressors. *Aggressive Behavior*, *43*, 263–272. doi:10.1002/ab.21686

de Bruyn, E. H., & Cillessen, A. H. N. (2006). Popularity in early adolescence: Prosocial and antisocial subtypes. *Journal of Adolescent Research*, *21*, 607–627. doi:10.1177/0743558406293966

Drummond, T. (1994). *The Student Risk Screening Scale (SRSS)*. Grants Pass, OR: Josephine County Mental Health Program.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Farmer, T. W., Farmer, E. M. Z., Estell, D. B., & Hutchins, B. (2007). The developmental dynamics of aggression and the prevention of school violence. *Journal of Emotional and Behavioral Disorders*, *15*, 197–208. doi:10.1177/10634266070150040201

Farmer, T. W., Hall, C. M., Leung, M.-C., Estell, D. B., & Brooks, D. (2011). Social prominence and the heterogeneity of

rejected status in late elementary school. *School Psychology Quarterly*, *26*, 260–274. doi:10.1037/a0025624

Farmer, T. W., Irvin, M. J., Motoca, L. M., Leung, M.-C., Hutchins, B. C., Brooks, D. S., & Hall, C. M. (2015). Externalizing and internalizing behavior problems, peer affiliations, and bullying involvement across the transition to middle school. *Journal of Emotional and Behavioral Disorders*, *23*, 3–16. doi:10.1177/1063426613491286

Forness, S. R., Freeman, S. F. N., Paparella, T., Kauffman, J. M., & Walker, H. M. (2012). Special education implications of point and cumulative prevalence for children with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *20*, 4–18. doi:10.1177/1063426611401624

Fröjd, S. A., Nissinen, E. S., Pelkonen, M. U., Marttunen, M. J., Koivisto, A.-M., & Kaltiala-Heino, R. (2008). Depression and school performance in middle adolescent boys and girls. *Journal of Adolescence*, *31*, 485–498. doi:10.1016/j.adolescence.2007.08.006

Green, J. G., Commer, J. S., Donaldson, A. R., Elkins, R. M., Nadeau, M. S., Reid, G., & Pincus, D. B. (2017). School functioning and use of school-based accommodations by treatment-seeking anxious children. *Journal of Emotional and Behavioral Disorders*, *25*, 220–232. doi:10.1177/1063426616664328

Holmes, B. W., Sheetz, A., Allison, M., Ancona, R., Attisha, E., Beers, N., . . . Young, T. (2016). Role of the school nurse in providing school health services. *Pediatrics*, *137*(6), e20160852. doi:10.1542/peds.2016-0852

Johnson, V., & Hutcherson, V. (2006). A study of the utilization patterns of an elementary school-based health clinic over a 5-year period. *Journal of School Health*, *76*, 373–378. doi:10.1111/j.1746-1561.2006.00128.x

Jolivette, K., Stichter, J. P., & McCormick, K. M. (2002). Making choices—Improving behavior—Engaging in learning. *TEACHING Exceptional Children*, *34*(3), 24–29. doi:10.1177/004005990203400303

Lane, K. L., Bruhn, A. L., Eisner, S. L., & Kalberg, J. R. (2010). Score reliability and validity of the Student Risk Screening Scale: A psychometrically sound, feasible tool for use in urban middle schools. *Journal of Emotional and Behavioral Disorders*, *18*, 211–224. doi:10.1177/1063426609349733

Lane, K. L., Kalberg, J. R., Parks, R. J., & Carter, E. W. (2008). Student Risks Screening Scale: Initial evidence for score reliability and validity at the high school level. *Journal of Emotional and Behavioral Disorders*, *16*, 178–190. doi:10.1177/1063426608314218

Lane, K. L., & Menzies, H. M. (2009). *Student Risk Screening Scale for Internalizing and Externalizing Behavior (SRSS-IE)*. Retrieved from http://www.ci3t.org/screening

Lane, K. L., Menzies, H. M., Bruhn, A. L., & Crnobori, M. E. (2011). *Managing challenging behaviors in schools: Research-based strategies that work*. New York, NY: Guilford Press.

Lane, K. L., Menzies, H. M., Oakes, W. P., Lambert, W., Cox, M. L., & Hankins, K. (2012). A validation of the Student Risk Screening Scale for internalizing and externalizing behaviors: Patterns in rural and urban elementary schools. *Behavioral Disorders*, *37*, 244–270.

Lane, K. L., Oakes, W. P., Cantwell, E. D., Menzies, H. M., Schatschneider, C., Lambert, W., & Common, E. A. (2017). Psychometric evidence of SRSS-IE scores in middle and high schools. *Journal of Emotional and Behavioral Disorders*, *25*, 233–245. doi:10.1177/1063426616670862

Lane, K. L., Oakes, W. P., Cantwell, E. D., Schatschneider, C., Menzies, H., Crittenden, M., & Messenger, M. (2016). Student Risk Screening Scale for internalizing and externalizing behaviors: Preliminary cut scores to support data-informed decision making in middle and high schools. *Behavior Disorders*, *42*, 271–284. doi:10.17988/BD-16-115.1

Lane, K. L., Oakes, W. P., Carter, E. W., Lambert, W., & Jenkins, A. (2013). Initial evidence for the reliability and validity of the Student Risk Screening Scale for internalizing and externalizing behaviors at the middle school level. *Assessment for Effective Intervention*, *39*, 24–38. doi:10.1177/1534508413489336

Lane, K. L., Oakes, W. P., Ennis, R. P., Cox, M. L., Schatschneider, C., & Lambert, W. (2013). Additional evidence for the reliability and validity of the Student Risk Screening Scale at the high school level: A replication and extension. *Journal of Emotional and Behavioral Disorders*, *21*, 97–115. doi:10.1177/1063426611407339

Lane, K. L., Oakes, W. P., Harris, P. J., Menzies, H. M., Cox, M. L., & Lambert, W. (2012). Initial evidence for the reliability and validity of the Student Risk Screening Scale for internalizing and externalizing behaviors at the elementary level. *Behavioral Disorders*, *37*, 99–122.

Lane, K. L., Oakes, W. P., & Menzies, H. M. (2014). Comprehensive, integrated, three-tiered models of prevention: Why does my school—and district—need an integrated approach to meet students' academic, behavioral, and social needs? *Preventing School Failure: Alternative Education for Children and Youth*, *58*, 121–128. doi:10.1080/1045988X.2014.893977

Lane, K. L., Oakes, W. P., Swogger, E., Schatschneider, C., Menzies, H. M., & Sanchez, J. (2015). Student Risk Screening Scale for Internalizing and Externalizing Behaviors: Preliminary cut scores to support data-informed decision making. *Behavioral Disorders*, *40*, 159–170. doi:10.17988/0198-7429-40.3.159

Lane, K. L., Parks, R. J., Kalberg, J. R., & Carter, E. W. (2007). Systematic screening at the middle school level: Score reliability and validity of the Students Risk Screening Scale. *Journal of Emotional and Behavioral Disorders*, *15*, 209–222. doi:10.1177/10634266070150040301

Lane, K. L., & Walker, H. M. (2015). The connection between assessment and intervention: How does screening lead to better interventions? In B. Bateman, M. Tankersley, & J. Lloyd (Eds.), *Enduring issues in special education: Personal perspectives* (pp. 283–301). New York, NY: Routledge.

Maggin, D. M., Wehby, J. H., Farmer, T. W., & Brooks, D. S. (2016). Intensive interventions for students with emotional and behavioral disorders: Issues, theory, and future directions. *Journal of Emotional and Behavioral Disorders*, *24*, 127–137. doi:10.1177/1063426616661498

Malecki, C. K., & Elliot, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal

analysis. *School Psychology Quarterly*, *17*, 1–23. doi:10.1521/scpq.17.1.1.19902

McIntosh, K., Campbell, A. L., Carter, D. R., & Zumbo, B. D. (2009). Concurrent validity of office discipline referrals and cut points used in schoolwide positive behavior supports. *Behavioral Disorders*, *34*, 100–113.

McIntosh, K., Frank, J. L., & Spaulding, S. A. (2010). Establishing research-based trajectories of office discipline referrals for individual students. *School Psychology Review*, *39*, 380–394.

McIntosh, K., & Goodman, S. (2015). *Integrating multi-tiered systems of support: Blending RTI and PBIS*. New York, NY: Guilford Press.

Menzies, H. M., Lane, K. L., Oakes, W. P., & Ennis, R. P. (2017). Increasing students' opportunities to respond: A strategy for supporting engagement. *Intervention in School and Clinic*, *52*, 204–209. doi:10.1177/1053451216659467

Moore, S. E., Norman, R. E., Suetani, S., Thomas, H. J., Sly, P. D., & Scott, J. G. (2017). Consequences of bullying victimization in childhood and adolescence: A systematic review and meta-analysis. *World Journal of Psychiatry*, *7*, 60–76. doi:10.5498/wjp.v7.i1.60

Oakes, W. P., Lane, K. L., Cox, M., & Messenger, M. (2014). Logistics of behavior screenings: How and why do we conduct behavior screenings at our school? *Preventing School Failure: Alternative Education for Children and Youth*, *58*, 183–190. doi:10.1080/1045988X.2014.895572

Rapport, M. D., Denney, C. B., Chung, K.-M., & Hustace, K. (2001). Internalizing behavior problems and scholastic achievement in children: Cognitive and behavioral pathways as mediators of outcome. *Journal of Clinical Child & Adolescent Psychology*, *30*, 536–551. doi:10.1207/s15374424jccp3004_10

Robertson, E. J., & Lane, K. L. (2007). Supporting middle school students with academic and behavioral concerns: A methodological illustration for conducting secondary interventions within three-tiered models of support. *Behavioral Disorders*, *33*, 5–22.

Rodkin, P. C. (2011). Bullying—And the power of peers. *Educational Leadership*, *69*, 10–16.

Shochet, I. M., Dadds, M. R., Ham, D., & Montague, R. (2006). School connectedness is an underemphasized parameter in adolescent mental health: Results of a community prediction study. *Journal of Clinical Child & Adolescent Psychology*, *35*, 170–179.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, *31*, 351–380. doi:10.1353/etc.0.0007

Siperstein, G. N., Wiley, A. L., & Forness, S. R. (2011). School context and the academic and behavioral progress of students with emotional disturbance. *Behavioral Disorders*, *36*, 172–184.

Vernberg, E. M., Nelson, T. D., Fonagy, P., & Twemlow, S. W. (2011). Victimization, aggression, and visits to the school nurse for somatic complaints, illnesses, and physical injuries. *Pediatrics*, *127*, 842–848. doi:10.1542/peds.2009-3415

Wagner, M. M. (1995). Outcomes for youths with serious emotional disturbance in secondary school and early adulthood. *Future of Children*, *5*, 90–112. doi:10.2307/1602359

Walker, H. M., Forness, S. R., & Lane, K. L. (2014). Design and management of scientific research in applied school settings. In B. Cook, M. Tankersley, & T. Landrum (Eds.), *Advances in learning and behavioral disabilities* (Vol. 27, pp. 141–169). Bingley, UK: Emerald.

Walker, H. M., Irvin, L., Noell, J., & Singer, G. (1992). A construct score approach to the assessment of social competence: Rationale, technological considerations, and anticipated outcomes. *Behavior Modification*, *16*, 448–474. doi:10.1177/01454455920164002

Walker, H. M., Severson, H. H., & Feil, E. G. (2014). *Systematic Screening for Behavior Disorders (SSBD) technical manual: Universal screening for preK–9* (2nd ed.). Eugene, OR: Pacific Northwest.

Watson, R. (2015, October). *Keynote address*. Kansas State Department of Education Annual Conference, Kansas City, KS.

Yudin, M. (2014, October). *PBIS: Providing opportunity*. Keynote address presented at the National Positive Behavioral Intervention and Supports Leadership Forum, PBIS Building Capacity & Partnerships to Enhance Educational Reform, Rosemont, IL.