Student Risk Screening Scale for Internalizing and Externalizing Behaviors: Preliminary Cut Scores to Support Data-Informed Decision Making

Kathleen Lynne Lane
University of Kansas

Wendy Peia Oakes
Arizona State University

Emily D. Swogger
University of Kansas

Christopher Schatschneider
Florida State University

Holly Mariah Menzies
California State University, Los Angeles

Jeremy Sanchez
Arizona State University

ABSTRACT: We report findings of a convergent validity study examining the internalizing subscale (SRSS-I5) of the newly adapted Student Risk Screening Scale for Internalizing and Externalizing (SRSS-IE12) with the internalizing subscale of the Teacher Report Form (TRF; Achenbach, 1991) conducted in 13 schools across three states with 195 kindergarten through fifth-grade elementary students. Results of logistic regression and receiver operating characteristics curves suggest the following cut scores for the SRSS-I5: 0–1 low risk, 2–3 moderate risk, and 4–15 high risk for internalizing behavior patterns, relative to the internalizing subscale of the TRF subscale. Limitations and future directions are discussed, along with illustrations of how to use the SRSS-I5 scores for decision making in a tiered system of supports.

■ The Student Risk Screening Scale (SRSS; Drummond, 1994) is a free-access, brief, systematic screening tool developed to detect elementary-age youth with antisocial tendencies. This seven-item tool requires approximately 10–15 min of teacher time to rate an entire class, making this a highly feasible tool for use in a tiered system of supports to detect students with behavior challenges for whom primary prevention efforts alone are insufficient. More specifically, using a 4-point Likert-type scale (*never* = 0, *occasionally* = 1, *sometimes* = 2, *frequently* = 3), teachers rate each student on seven items: (a) steal; (b) lie, cheat, sneak; (c) behavior problem; (d) peer rejection; (e) low academic achievement; (f) negative attitude; and (g) aggressive behavior. These items are summed to obtain a total score for each student and used to place students into one of the following risk categories developed by Drummond (1994): low (0–3), moderate (4–8), or high (9–21).

These data can be used for a number of purposes within of the framework of tiered systems of support as well as in schools that do not yet employ graduated systems of support. For example, screening data can be analyzed to examine the overall level of risk in a building or district, ascertain if teacher-level strategies and supports (e.g., increasing opportunities to respond, incorporating instructional choice) may be warranted, and determine students in need of validated Tier 2 (for some) and Tier 3 (for

a few) supports (Lane & Walker, 2015; Oakes, Lane, Cox, & Messenger, 2014).

While the SRSS can be completed in a paper-and-pencil format, in the last 10 years many school districts and state technical assistance projects (e.g., see Michigan's Integrated Behavior and Learning and Support Initiative [MiBLSi.cenmi.org] and the Pennsylvania Department of Education [PaTTAN.net] websites) have developed free-access electronic structures. The easy and affordable structures and resources have enabled school systems to incorporate this screening measure as part of regular, district-wide school practices.

Feasible and Effective

Screening tools are considered feasible if they fit within the resources of the district or school site and provide useful information to educators for intervening at the earliest sign of need (Oakes et al., 2014). The SRSS addresses issues of feasibility along a number of dimensions such as time, cost, and ease of use. For example, when considering district-level personnel's time to prepare and manage data structures, teacher's time to complete the screener three times per year, and teacher and leadership team members' time for scoring screening tools, the overall amount of time involved when using the SRSS is quite nominal. Furthermore, the associated costs for the SRSS are low relative to the costs of purchasing a new commercially available system coupled with the needed professional learning to use such a system. District-developed systems to use the SRSS often involve using Excel spreadsheets kept on secure teacher drives managed by district technology specialists and completed by school-site personnel during regularly scheduled faculty meetings. Such systems often require a minimal amount of training for teachers and district administrators to use the screening measure with fidelity to ensure accurate decisions regarding tiered intervention supports.

In addition to being feasible, the SRSS is effective. A series of psychometric studies conducted by Drummond and colleagues reported SRSS total scores predicted undesirable behavioral and academic outcomes 1.5 to 10 years following initial nominations (Drummond, Eddy, Reid, & Bank, 1994). The earlier work by Drummond (1994) also established initial evidence of convergent validity between SRSS scores and the Child Behavior Checklist's aggressive behavior subscale score (Achenbach, 1991). More recently, Lane, Menzies, and colleagues (2012) conducted a series of psychometric studies indicating SRSS scores as used across the K–12 continuum have demonstrated internal consistency ($>.80$) and test–retest score stability, broadening the uses of SRSS to middle and high school students.

At the elementary level, SRSS scores predicted year-end behavioral performance as measured by office discipline referrals (ODRs) earned over the course of the academic year (Menzies & Lane, 2012; Oakes et al., 2010) and self-control skills as measured by the Social Skills Rating System subscale scores (Gresham & Elliott, 1990; Menzies & Lane, 2012). In addition, SRSS scores predicted elementary-age students' year-end reading performance (Menzies & Lane, 2012; Oakes et al., 2010). In examining psychometric rigor of SRSS scores relative to the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1992), which is often referred to as the gold standard of systematic screening (Kauffman & Brigham, 2009), SRSS scores were equally sensitive and specific in identifying students with externalizing (improving chance estimates by approximately 45%) and, to a lesser extent, those with internalizing behaviors (improving chance estimates approximately 30%; Lane, Kalberg, Lambert, Crnobori, & Bruhn, 2010; Lane et al., 2009).

Focusing on Detection of Internalizing Behaviors

Although the SRSS was not developed to identify students with internalizing behaviors (e.g., shy, anxious, socially withdrawn), psychometric studies conducted to date suggest SRSS scores hold promise for detecting students with these more veiled or covert behavior challenges. This is important given that recent point prevalence estimates indicate upwards of 20% of students have at least some mild manifestations of emotional or behavioral disorders (EBD), including both externalizing and internalizing behavior patterns (Forness, Freeman, Paparella, Kauffman, & Walker, 2012). Given the more subtle nature of their behavior challenges, students with internalizing behaviors are less likely to be identified and receive assistance compared to students with externalizing behaviors (Bradshaw, Buckley, & Ialongo, 2008). This is troublesome as

internalizing behaviors are more prevalent than many people realize, negatively impact students' educational experiences, and often continue into adulthood (Breton et al., 1999; Costello, Erkanli, & Angold, 2006; Malecki & Elliott, 2002; Rapport, Denny, Chung, & Hustace, 2001). As such, it is important to develop feasible, effective, and efficient approaches for detecting students with internalizing behaviors who are often overlooked by examining data more common to school practices (e.g., ODRs; McIntosh, Campbell, Carter, & Zumbo, 2009).

To capitalize on the ease with which the SRSS can be completed and the increasingly wide-scale adoption of the SRSS, Lane, Menzies, and colleagues (2012) developed an adapted measure of the SRSS by exploring the utility of seven additional items characteristic of internalizing behaviors to expand the scope of this tool to detect students with internalizing as well as externalizing behavior patterns. In the first psychometric study of this adapted tool, Student Risk Screening Scale: Internalizing and Externalizing (SRSS-IE), Lane, Oakes et al. (2012) examined the reliability of the SRSS-IE14 scores for use at the elementary level (N = 2,460 students) using a data analytic plan grounded in Classical Test Theory (Cronbach & Shavelson, 2004). They examined item level data, computed internal consistency estimates, and explored the factor structure of the SRSS-IE14. Results provided preliminary evidence to support retaining five of the initially proposed seven items– (a) emotionally flat; (b) shy, withdrawn; (c) sad; depressed; (d) anxious; and (e) lonely; leading to the SRSS-IE12. They also reported convergent validity between SRSS-IE12 scores with SSBD scores (Walker & Severson, 1992) as well as Strength and Difficulties Questionnaire scores (SDQ; Goodman, 1997), the latter of which is another free access systematic screener, although more time consuming to complete and score.

Next, Lane, Menzies et al. (2012) conducted another set of studies to continue to explore the generalizability of the initial findings. In this article, they reported outcomes from one study conducted in a rural school district (N = 982; three schools), and a second conducted in an urban district (N = 1,079; four schools). Results of both studies, including item level, internal consistency, factor structure, and test–retest score stability analyses, supported retention of the same five items noted in the Lane, Oakes et al. (2012) study. Consistent with the initial findings offered by Lane, Oakes et al. (2012), results indicated the SRSS-IE12, as used with elementary students, is a two-factor structure. It yielded two subscales: SRSS-E7 (the original seven items constituting the SRSS) and the SRSS-I5 (the five retained items proposed to measure internalizing behaviors). Lane, Menzies et al. (2012) also reported convergent validity of the SRSS-IE12 scores, as well as the two subscales scores (SRSS-E7 and SRSS-I5), with the SSBD.

Yet, a key limitation of the study reported by Lane, Oakes et al. (2012) and the two studies reported by Lane, Menzies et al. (2012) is the absence of cut scores for the SRSS-I5. This is a critical next step as cut scores are used for decision making for students who may require Tier 2 and Tier 3 supports, informing teacher level interventions, and examining the overall level of internalizing risk in a school or district (Lane & Walker, 2015; Oakes et al., 2014). Specifically, these cut scores are needed for school-site teams to make explicit, transparent decisions to inform instructional programming. For example, a student who scores in the moderate-risk range for internalizing behaviors will likely have different intervention needs than a student who scores in the high-risk range for internalizing behaviors. Also, a large percentage of the student body scoring in the moderate- and/or high-risk range may have implications for schoolwide Tier 1 supports.

Purpose

The intent of this study is to address this void and offer preliminary cut scores for schools to use in decision making within—as well as independent of—tiered systems of support. Namely, this study is the first conducted to determine specific cut scores needed to identify students at low, moderate, and high risk for internalizing behaviors similar to those developed by Drummond for use in interpreting the SRSS (now referred to as SRSS-E7) scores.

This study is an important, yet modest, step in our programmatic line of inquiry to develop and validate the SRSS-IE for use in elementary, middle, and high schools. Here we focus specifically on the SRSS-I5 subscale as used with elementary-age students, with a goal of determining preliminary cut scores for this subscale. Future studies will be needed to determine the generalizability of the findings

TABLE 1
Student Characteristics

	State			
Variable/Level	**State A**	**State B**	**State C**	**Total**
Students	*n* = 99	*n* = 64	*n* = 32	*N* = 195
Gender, % (*n*)				
Male	55.56 (55)	56.25 (36)	56.25 (18)	55.90 (109)
Female	44.44 (44)	43.75 (28)	43.75 (14)	44.10 (86)
Ethnicity, % (*n*)				
White	80.81 (80)	46.88 (30)	3.13 (1)	56.92 (111)
Black	8.08 (8)	7.81 (5)	6.25 (2)	7.69 (15)
Hispanic	2.02 (2)	40.63 (26)	84.38 (27)	28.21 (55)
Asian	2.02 (2)	0.00 (0)	6.25 (2)	2.05 (4)
Mixed races	7.07 (7)	4.69 (3)	0.00 (0)	5.13 (10)
Grade level, % (*n*)				
Kindergarten	10.10 (10)	7.81 (5)	9.38 (3)	9.23 (18)
First	16.16 (16)	14.06 (9)	25.00 (8)	16.92 (33)
Second	16.16 (16)	14.06 (9)	3.13 (1)	13.33 (26)
Third	17.17 (17)	23.44 (15)	43.75 (14)	23.59 (46)
Fourth	20.20 (20)	15.63 (10)	6.25 (2)	16.41 (32)
Fifth	20.20 (20)	25.00 (16)	12.50 (4)	20.51 (40)
Age, *M* (*SD*)	9.10 (1.60)	—	8.55 (1.51)	

Note. Age was not provided for students in State B. Dash refers to data not available.

reported here. This initial evidence must be supported through future research as school-site leadership teams who are seeking a tool that offers a balance between psychometric soundness and feasibility will use the SRSS-I5 to detect elementary students with different facets of EBD who may require Tier 2 and Tier 3 supports.

We conducted this current validation study to determine the reliability and validity of a set of five items designed to extend the scope of the SRSS to detect elementary-age students with internalizing behavior patterns. Specifically, in this paper we examined the convergent validity of these five items (SRSS-I5) with the internalizing subscale of the Teacher Report Form (TRF; Achenbach, 1991) rating scale to establish preliminary cut scores to inform decision making regarding instructional programming for elementary students.

Method

Participants and Setting

Participants were 195 kindergarten through fifth grade elementary students (109 boys, 86 girls) attending one of 13 elementary schools across three states (five in Missouri, five in Kansas, and three in Arizona; see *Table 1* for Student Characteristics and *Table 2* for School Characteristics). The majority of students were White (56.92%), followed by 28.21% Hispanic, 7.69% Black, 2.05% Asian, and 5.13% were of mixed race. None of the students were receiving special education services.

Procedures

After securing university and district approvals for two projects (one conducted in Missouri and Arizona during the 2012–2013 academic year and a second conducted in Kansas during the 2013–2014 academic year), all kindergarten through fifth grade teachers were invited to participate from the 13 elementary schools where behavior screenings were taking place as part of regular school practices. An introductory meeting was held by the principal investigators with teachers from all schools to explain the purpose of the study, answer questions, and invite teachers to participate. Teacher consent forms were dis-

TABLE 2
School Characteristics

State	School	Attendance Rate[a] %	Classroom Teachers[b] *N*	Enrollment *N*	FRPL %	Locale	AMO Status[c]	Title 1 Eligible
State A	1	97.0	26.27	425	10.1	Suburb: Large	Not met	Yes
	2	95.7	23.96	379	12.7	Suburb: Large	Met	Yes
	3	96.5	25.62	426	22.5	Suburb: Large	Not met	Yes
	4	96.4	32.20	559	6.6	Suburb: Large	Not met	No
	5	97.7	26.10	429	6.8	Suburb: Large	Met	No
State B	1	95.8	25.47	368	67.9	City: Midsize	Met	Yes
	2	95.9	36.68	393	95.2	City: Midsize	Met	Yes
	3	94.6	32.01	356	91.6	City: Midsize	Not met	Yes
	4	95.2	35.93	442	86.2	City: Midsize	Not met	Yes
	5	95.5	53.56	604	87.7	City: Midsize	Met	Yes
State C	1	—	44.00	962	99.7	City: Large	Not met	Yes
	2	—	49.00	1,095	49.1	Suburb: Small	Not met	Yes
	3	—	32.80	656	99.5	Rural: Fringe	Not met	Yes

Note. Sources: National Center for Education Statistics, Common Core Data 2011–2012. FRPL = free and reduced-price lunch eligible; AMO = Annual Measurable Objectives. Dash refers to data not available.
[a]2012–2013 school report card data.
[b]Reported as full time equivalents.
[c]The No Child Left Behind Act (1997) established improvement goals measured by AMO in Reading/Communication Arts and Math. For State A, if school did not meet all subscales in both content areas, AMO were not met. For State B, if subscale for growth was met, AMO was considered to be met. State C, AMO met or not met status was reported. State C school information is reported for students K–8.

tributed and collected at the end of each meeting. One return trip was made to collect any remaining consent forms. See *Table 3* for the titration of the consenting process, resulting in the 211 teachers agreeing to participate in this study by completing the TRF for up to four students in their class.

Behavior screenings were being introduced at each school as part of regular school practices, with a goal of using these data to inform instruction and connect students with Tier 2 and Tier 3 supports as part of their respective multitiered system of supports. We analyzed de-identified student level data from the Student Risk Screening Scale for Internalizing and Externalizing Behavior (SRSS-IE; Lane & Menzies, 2009) winter administrations (late November to early December) for each teacher's class to identify four categories of students' risk: (a) no risk: zero ratings on all SRSS-IE items; (b) externalizing only: moderate or high risk on the original seven items constituting the SRSS (SRSS-E7 total score of 4–21) and no internalizing concerns according to the newly added five items (SRSS-I5 total score of 0); (c) internalizing only: SRSS-E7 total score of zero and SRSS-I5 score of 3–15 (arbitrarily decided to obtain some behavioral risk on this dimension); and (d) combined concerns: SRSS-E7 total scores of 4–21 and SRSS-I5 total score of 3–12). We randomly selected one student from each category to obtain parent consent.

We prepared invitation letters for teachers with the randomly selected student identification numbers in the four categories. We also provided teachers with a second randomly selected student in each category in case the first student's parents declined. Teachers received up to eight envelopes with prepared parent consent forms (two copies). Teachers wrote ''to the parents of XXX,'' as we did not have any student names, nor any method of identifying students prior to receiving consent. The envelopes contained two copies of the parent consent, which requested parents to discuss this project with their child as they deemed appropriate. Parents were informed that the information from the TRF was to be used only for research purposes and would not be shared with the school or the parents. While there are some items on the screening tools where teachers report knowledge of extreme behavior problems, teachers would

TABLE 3
Consenting Process

Phase	School	State			
		State A	State B	State C	Total
Teachers invited, *n*		125	98	66	289
	1	23	16	25	
	2	26	19	24	
	3	23	15	17	
	4	28	22	—	
	5	25	26	—	
Participating teachers, % (*n*)		83.20 (104)	74.49 (73)	68.18 (45)	76.82 (222)
	1	91.30 (21)	75.00 (12)	72.00 (18)	
	2	65.38 (17)	63.16 (12)	58.33 (14)	
	3	95.65 (22)	60.00 (9)	76.47 (13)	
	4	82.14 (23)	100.00 (22)	—	
	5	84.00 (21)	69.23 (18)	—	
Students invited in participating teachers' classes, *n*		272	209	82	563
	1	48	31	47	
	2	50	36	16	
	3	72	39	19	
	4	47	71	—	
	5	55	32	—	
Parental consent obtained for students invited in participating teachers' classes, % (*n*)		38.60 (105)	33.97 (71)	42.68 (35)	37.48 (211)
	1	37.50 (18)	45.16 (14)	51.06 (24)	
	2	52.00 (26)	25.00 (9)	18.75 (3)	
	3	41.67 (30)	28.21 (11)	42.11 (8)	
	4	23.40 (11)	35.21 (25)	—	
	5	38.36 (20)	37.50 (12)	—	
TRFs completed for students for whom parents provided consent, % (*n*)		100.00 (105)	100.00 (71)	100.00 (35)	100.00 (211)
	1	100.00 (18)	100.00 (14)	100.00 (24)	
	2	100.00 (26)	100.00 (9)	100.00 (3)	
	3	100.00 (30)	100.00 (11)	100.00 (8)	
	4	100.00 (11)	100.00 (25)	—	
	5	100.00 (20)	100.00 (12)	—	
Teachers completing TRFs, % (*n*)		51.92 (54)	60.27 (44)	28.89 (13)	50.00 (111)
	1	52.38 (11)	58.33 (7)	50.00 (9)	
	2	70.59 (12)	50.00 (6)	14.29 (2)	
	3	59.09 (13)	66.67 (6)	15.38 (2)	
	4	34.78 (8)	77.27 (17)	—	
	5	47.62 (10)	44.44 (8)	—	

Note. TRF = Teacher Report Form (Achenbach, 1991). Differences in the *n* for teachers participating and teachers completing TRFs were due to teacher withdrawal ($n = 2$) or no return of parent consents. Differences in the *n* for students for whom parents provided consent and the *n* reported in *Table 1* of the final sample were due to TRFs completed for students younger than 6 years old. Dash refers to data not available.

have already reported this information under mandated reporting requirements.

Teachers who agreed to participate completed a brief teacher demographic information form (e.g., gender, years of experience, highest degree obtained) requiring about 10 min of their time, as well as the full TRF (description to follow) for up to four students (approximately 20 min of teacher time for each student, or up to 1 hr, 20 min total) from their class after parent permission was secured. Teachers completed a maximum of four TRFs (see *Table 3*) and returned them to research staff with parent consent forms.

After scoring the TRFs, data were entered and 25% checked for reliability of entry. To thank teachers for their participation, districts received a research brief with information about systematic screening and preliminary findings. In addition, principal investigators offered to work with teachers (e.g., 2-hr sessions) to teach school faculty how to construct intervention grids to connect students to evidence-based practices to meet academic, behavioral, and social needs as determined by analyzing behavior screening data in conjunction with other data collected as part of regular school practices (e.g., academic measures, attendance, and ODRs).

Measures

Student Risk Screening Scale—Internalizing and Externalizing (SRSS-IE12)

The SRSS-IE12 is an adapted version of the SRSS designed for use at the elementary level (Drummond, 1994). The SRSS-IE includes the original seven items as well as an additional five items characteristic of internalizing behaviors retained based on results of a data analytic plan grounded in Classical Test Theory (Lane, Menzies et al., 2012; Lane, Oakes et al., 2012). The SRSS-IE12 contains two subscale scores. The SRSS-E7 includes the original seven items, retained without adaptation: (a) steal; (b) lie, cheat, sneak; (c) behavior problems; (d) peer rejection; (e) low academic achievement; (f) negative attitude; and (g) aggressive behavior. The SRSS-I5 includes: (a) emotionally flat; (b) shy, withdrawn; (c) sad, depressed; (d) anxious; and (e) lonely. Teachers rate all 12 items using the 4-point, Likert-type scale developed by Drummond (1994) as follows: 0 = *never*, 1 = *occasionally*, 2 = *sometimes*, and 3 = *frequently*.

Child Behavior Checklist: Teacher Report Form

The TRF (Achenbach, 1991) is part of the Achenbach System of Empirically Based Assessment (ASEBA) and is a widely used behavior rating scale, developed to detect both internalizing and externalizing behavioral dimensions. In addition to the 10 demographic items, teachers complete the 113 items constituting the TRF for children ages 6 to 18 using a 3-point Likert-type scale: 0 = *not true (as far as you know)*, 1 = *somewhat or sometimes true*, and 2 = *very true or often true within the past two months*. The broadband disorders contained the following empirically-based syndrome scales: Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problem, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior, and the following DSM-oriented scales: Affective Problems, Anxiety Problems, Somatic Complaints, Attention Deficit/Hyperactivity Problems (Inattention and Hyperactivity–Impulsivity), Oppositional Defiant Problems, and Conduct Problems. The TRF has extensive evidence for its use in identifying these disorders. Test–retest score reliability was $r = .90$, with alpha coefficients of .90 internalizing, .97 externalizing, and .97 total problems. Internal consistency alpha coefficients for the empirically based syndrome scales ranged from .72 to .95, and for the DSM scales were reported at .73 to .94 (Achenbach & Rescorla, 2001).

Data Analysis Plan

The goal of the data analytic plan was to explore convergent validity (Campbell & Fiske, 1959; Fiske & Campbell, 1992) between the SRSS-I5 and the internalizing broadband scores of the TRF to develop two cut scores on the SRSS-I5 that would correspond to our best prediction of students who were experiencing no internalizing problems, borderline problems, or clinical levels of internalizing problems. To obtain these estimates, we constructed two logistic regression models along with two receiver operator characteristic (ROC) curves to examine the accuracy (sensitivity and specificity) of the SRSS-I5 compared to the internalizing TRF scores (Petras, Chilcoat, Leaf, Ialongo, & Kellam, 2004). The first logistic regression model predicted students who would be classified as having no internalizing problems versus those having either

borderline or clinical levels of internalizing problems on the TRF. A second logistic regression model was fit to predict students who were displaying clinical levels of internalizing problems from those who were not on the TRF. In both models, the area under the curve (AUC) ranges from .50 (a chance estimate) to 1.0 (perfect prediction), with scores higher than .50 and lower than 1.0 indicating better-than-chance predictions (Bunevicius, Peceliuniene, Mickuviene, Valius, & Bunevicius, 2007). Essentially, ROC curves can be conceptualized as an array of possible cutting scores, with each cut score offering a unique balance of benefit (sensitivity) and cost (specificity; Lane et al., 2009).

Results

SRSS-I5 Cut Score #1: Low Internalizing versus Borderline or Clinical Levels

The first logistic regression model predicted students classified as having no internalizing problems versus those having either borderline or clinical levels of internalizing problems on the TRF. The SRSS-I5 scale significantly predicted group membership (logit estimate of .41, chi-square = 24.7, $p < .0001$, odds ratio = 1.5).

To select a cut score on the SRSS-I5, we fit an empirical ROC curve, and selected the place on the curve that would yield an optimal level of sensitivity and specificity. Sensitivity is a number that ranges from 0 to 1.0 and represents the percentage of students who are rated as having a problem on the TRF that are detected by the SRSS-I5 screening tool. Specificity has the same range as sensitivity, and represents the percentage of students who are not rated as having internalizing problems that are rated by the screening tool as not having a problem. A visual inspection of the ROC curve data indicated that a cut score of less than 2 yields a sensitivity of .70 and a specificity of .83, and an overall correct classification rate of .81. Thus, a SRSS-I5 total score of 0–1 was designated as low risk for internalizing problems.

SRSS-I5 Cut Score #2: Low or Borderline Internalizing versus Clinical Levels

A second logistic regression model was fit to predict students who were displaying clinical levels of internalizing problems from those who were not. The SRSS-I5 scale significantly predicted group membership for those students exhibiting clinical levels of internalizing problems (logit estimate of .44, chi-square = 25.5, $p < .0001$, odds ratio =1.6).

To select the second cut score on the SRSS-I5, we fit an empirical ROC curve, and selected the place on the curve that would again yield an optimal level of sensitivity and specificity. A visual inspection of the ROC curve information indicated that a cut score of 4 or greater on the SRSS-I5 yields a sensitivity of .78 and a specificity of .81, and an overall correct classification rate of .81. Thus, a SRSS-I5 total score of 4–15 was designated as high risk for internalizing behavior, leaving a SRSS-I5 total score of 2–3 designated as moderate risk.

In sum, results from the current study suggest cut scores for the SRSS-I5 when examining internalizing behavior of 0–1 for low risk, 2–3 for moderate risk, and 4–15 for high risk yield groupings with acceptable levels of sensitivity and specificity.

Discussion

Students with internalizing behaviors are often undetected and consequently not supported within the context of tiered systems of supports given the often covert nature of their behaviors (Bradshaw et al., 2008; Forness et al., 2012). While internalizing behavior patterns are far more veiled than externalizing behaviors, which are often quick to capture teacher attention, they are no less serious (Costello et al., 2006). It is critical for the research and teaching communities to collaborate to develop and test feasible and effective screening tools to detect students with internalizing behaviors, as current methods of detection such as relying on ODR data are not effective with this particular group of students (McIntosh et al., 2009).

We are pleased to offer findings from this modest, yet important, study of the SRSS-IE to determine cut scores for the internalizing subscale, the SRSS-I5, intended to support decision making within tiered systems of support. Results of logistic regression models predicting who would be classified on the TRF as having (a) no internalizing problems versus those having either borderline or clinical levels of internalizing (logit estimate of .41, chi-square = 24.7, $p < .0001$, odds ratio = 1.5), and (b) clinical level of internalizing problems

versus those who were not (logit estimate of .44, chi-square = 25.5, $p < .0001$, odds ratio =1.6) were both statistically significant. Visual inspections of ROC curve data indicated that cut scores of (a) less than 2 offered a sensitivity of .70 and a specificity of .83, and an overall correct classification rate of .81, and (b) 4 or more offered a sensitivity of .78 and a specificity of .81, and an overall correct classification rate of .81. While preliminary in nature, findings from this study yield the following proposed cut scores for the SRSS-I5: 0–1 suggest low risk, 2–3 suggest moderate risk, and 4–15 suggest high risk for internalizing behavior patterns.

This study extends the knowledge base of the utility of the SRSS-IE. Results of previous studies established convergent validity of the SRSS-IE12 with other validated screening tools: the SSBD and the SDQ (Lane, Menzies et al., 2012; Lane, Oakes et al., 2012), whereas this study compares the SRSS-IE screening tool (specifically, the SRSS-I5 score) to a well-established behavior rating scale—the TRF internalizing broadband score. Rather than having to complete the 113 items on the TRF (a widely used and psychometrically sound rating scale), teachers can spend far less time completing the five items constituting the SRSS-I5 subscale score to predict which elementary-age students have internalizing behaviors.

To illustrate, teachers might begin to use these cut scores in the same fashion as the original SRSS (referred to as the SRSS-E7) cut scores: (a) to examine the overall level of risk in a school, (b) to explore risk in a given grade level or class, and (c) to determine which students may require Tier 2 (moderate risk) and Tier 3 (high risk) interventions and supports. In the latter example, a student with an SRSS-I5 score of 3 would suggest moderate risk for internalizing behavior challenges. Before moving forward in connecting this student with a Tier 2 support specific to this internalizing concern, the school-site leadership team would first look at the overall level of internalizing issues in a school by examining the SRSS-I5 scores to determine the percentage of the student body placed into low (0–1), moderate (2–3), or high (4–15) categories. If 80% of the student body is scoring in the low risk range, then it is does not suggest a system-level concern (which may call for addressing primary prevention efforts). Next, within a multitiered system of supports they might look at the homeroom teachers' treatment integrity scores for their Tier 1 efforts to make certain all students in that class are accessing Tier 1 practices (e.g., Germer et al., 2011). In addition, students' attendance data are considered to make certain that individual students are present at school on a regular enough basis to access Tier 1 efforts (Lane, Oakes, Ennis, & Hirsch, 2014). Next, the teacher may look at his or her individual class ratings to determine the percentage of students in his or her class placing into the three risk categories. If the data suggest 30% or more of students are placing into the moderate risk range according to SRSS-I5 scores, then the teacher might begin with a classwide intervention (e.g., increasing opportunities to respond or use of instructional choice; Lane, Menzies, Ennis, & Oakes, in press). Yet, if Tier 1 is in place, the student is present, and there are not classwide concerns, then it would be appropriate to analyze these data in conjunction with other data (e.g., academic screening data) to connect the student with evidenced-based strategies for internalizing behaviors (see Kern, Hilt-Panahon, & Mukherjee, 2013).

Limitations and Future Directions

In the previous illustration and in actual school contexts, we recommend caution when interpreting the proposed cut scores. We offer these scores with cautious optimism, respectfully requesting these data be used sagaciously and in conjunction with other school data—interpreted in light of the following limitations.

First, as with any screening tool used in a low base rate context, this measure (the sum of the five proposed internalizing items, the SRSS-I5 score) will likely identify some students who are not exhibiting internalizing behaviors (false positives; Walker, Forness, & Lane, 2014). Since the majority of students do not have an internalizing disorder, most classification errors when using the screening score will falsely declare that a student has higher than average internalizing behaviors, when in fact they do not. This could potentially be a problem if the result of being detected by the screening tool carries high-stakes consequences. However, if being identified as at risk for internalizing behaviors according to the screening score has low-stakes consequences, such as increased monitoring or low-intensity supports, these scores will likely be more useful as they

involve minimal effort and will most likely be helpful to the student.

Second, this study does not address the nested nature of the data, meaning that teachers were rating up to four students in their classrooms (one in each of the four categories noted in the method). The result of ignoring the nested structure of the data is that the standard errors associated with the odds ratios will be smaller than would be expected given the assumptions of the statistic. This violation of the assumption of independence of observation is not expected to impact the estimated size of the odds ratio, but could impact the test of statistical significance. We encourage future inquiry with larger samples of elementary-age students to explore this possible clustering effect.

Third, to establish the generalizability of these findings at the elementary level, replication is critical. While this study includes elementary students from three states representing 13 schools and three districts, it is necessary to replicate this study in other locales that also serve diverse populations.

Similarly, it will be important to conduct future inquiry to explore the use of the SRSS-IE in middle and high school settings to detect and subsequently support adolescents with internalizing behavior patterns. Exploring the use of the SRSS-IE with older students is important given the transition from elementary to middle and middle to high school can be formidable even for the most capable students, and especially those at risk for EBD (Akos & Galassi, 2004; Lane, Oakes, Carter, & Messenger, 2015).

Summary

We respectfully submit this modest—but important—study as part of our programmatic line of inquiry to develop and validate this adapted tool (SRSS-IE) for use in elementary, middle, and high schools. In this study we examined the SRSS-I5 internalizing subscale as used with elementary-age students, with the intent of providing research and teaching communities with preliminary cut scores for this subscale. We encourage people to use these scores cautiously as part of decision making within tiered systems of supports. We also encourage other research teams to conduct future studies to replicate this study at the elementary level and begin to explore cut scores with middle and high schools, with a goal of providing a free-access, feasible, effective, efficient screening tool for use across the K–12 continuum as part of tiered systems of support.

REFERENCES

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

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

Akos, P., & Galassi, J. P. (2004). Middle and high school transitions as viewed by students, parents, and teachers. *Professional School Counseling, 7*, 212–221.

Bradshaw, C. P., Buckley, J., & Ialongo, N. (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

Breton, J. J., Bergeron, L., Valla, J. P., Berthiaume, C., Gaudet, N., Lambert, J., . . . Lepine, S. (1999). Quebec child mental health survey: Prevalence of SDM-III-R mental health disorders. *Journal of Child Psychology and Psychiatry, 40*, 375–384.

Bunevicius, A., Peceliuniene, J., Mickuviene, N., Valius, L., & Bunevicius, R. (2007). Screening for depression and anxiety disorders in primary care patients. *Depression and Anxiety, 24*, 455–460. doi:10.1002/da.20274

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. doi:10.1037/h0046016

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry, 47*, 1263–1271.

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391–418. doi:10.1177/0013164404266386

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

Drummond, T., Eddy, J. M., Reid, J. B., & Bank, L. (November, 1994). *The Student Risk Screening Scale: A brief teacher screening instrument for conduct disorder*. Paper presented at the Fourth Annual Prevention Conference, Washington, DC.

Fiske, D. W., & Campbell, D. T. (1992). Citations do not solve problems. *Psychological Bulletin, 112*, 393–395. doi:10.1037//0033-2909.112.3.393

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

Germer, K. A., Kaplan, L. M., Giroux, L. N., Markham, E. H., Ferris, G., Oakes, W., & Lane, K. L. (2011). A function-based intervention to increase a second-grade student's on-task behavior in a general education classroom. *Beyond Behavior, 20*, 19–30.

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

Gresham, F. M., & Elliott, S. N. (1990). *The Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Kauffman, J. M., & Brigham, F. J. (2009). *Working with troubled children*. Verona, WI: Full Court Press.

Kern, L., Hilt-Panahon, A., & Mukherjee, A. A. (2013). Strategies to address internalizing behavior problems. In K. L. Lane, B. G. Cook, & M. Tankersley (Eds.), *Research-based strategies for improving outcomes in behavior* (pp. 59–72). New York: Pearson.

Lane, K. L, Kalberg, J. R., Lambert, W., Crnobori, M., & Bruhn, A. (2010). A comparison of systematic screening tools for emotional and behavioral disorders: A replication. *Journal of Emotional and Behavioral Disorders, 18*, 100–112. doi:10.1177/1063426609341069

Lane, K. L., Little, M. A., Casey, A. M., Lambert, W., Wehby, J. H., Weisenbach, J. L., & Phillips, A. (2009). A comparison of systematic screening tools for emotional and behavioral disorders: How do they compare? *Journal of Emotional and Behavioral Disorders, 17*, 93–105.

Lane, K. L., & Menzies, H. M. (2009). *Student Risk Screening Scale for Early Internalizing and Externalizing Behavior (SRSS-IE)*. Unpublished rating scale.

Lane, K. L., Menzies, H. M., Ennis, R. P., & Oakes, W. P. (in press). *Supporting behavior for school success: A step-by-step guide to key strategies*. 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., Carter, E., W., & Messenger, M. (2015). Examining behavioral risk and academic performance for students' transition from elementary to middle school. *Journal of Positive Behavior Interventions and Support, 17*, 39–49. doi:10.1177/1098300714524825

Lane, K. L., Oakes, W. P., Ennis, R. P., & Hirsch, S. E. (2014). Identifying students for secondary and tertiary prevention efforts: How do we determine which students have Tier 2 and Tier 3 needs? *Preventing School Failure, 58*, 171–182. doi:10.1080/1045988X.2014.895573

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., & Walker, H. M. (2015). The connection between assessment and intervention: How does screening lead to better interventions? In B. Bateman, J. W. Lloyd & M. Tankersley (Eds.), *Enduring issues in special education: Personal perspectives* (pp. 285–301). New York, NY: Routledge.

Malecki, C. K., & Elliott, 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 support. *Behavioral Disorders, 34*, 100–113.

Menzies, H. M., & Lane, K. L. (2012). Validity of the student risk screening scale: Evidence of predictive validity in a diverse, suburban elementary setting. *Journal of Emotional and Behavioral Disorders, 20*, 82–91. doi:10.1177/1063426610389613

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, 58*, 159–170. doi:10.1080/1045988X.2014.895572

Oakes, W. P., Wilder, K., Lane, K. L., Powers, L., Yokoyama, L., O'Hare, M. E., & Jenkins, A. B. (2010). Psychometric properties of the Student Risk Screening Scale: An effective tool for use in diverse urban elementary schools. *Assessment for Effective Intervention, 35*, 231–239. doi:10.1177/1534508410379796

Petras, H., Chilcoat, H., Leaf, P. J., Ialongo, N. S., & Kellam, S. G. (2004). Utility of TOCA-R scores during the elementary school years in identifying later violence among adolescent males. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 88–96. doi:10.1097/00004583-200401000-00018

Rapport, M. D., Denny, C. B., Chung, K., & Hustace, K. (2001). Internalizing behavior problems and scholastic achievement in children: Cognitive mediators and pathways as mediators of outcomes. *Journal of Clinical Child Psychology, 30*, 536–551. doi:10.1207/S15374424JCCP3004_10

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., & Severson, H. (1992). *Systematic Screening for Behavior Disorders: Technical manual*. Longmont, CO: Sopris West.

AUTHORS' NOTE

This project was funded in part by a School of Education Research Grant from The University of Kansas: Identifying and supporting K–12 students within the context of three-tiered models of prevention to meet students' multiple needs: A collaborative effort (2013–2014).

Address correspondence to Kathleen Lynne Lane, Professor, University of Kansas, Department of Special Education, 1122 West Campus Road, JRP Room 720, Lawrence, KS 66045; E-mail: Kathleen.Lane@ku.edu.

MANUSCRIPT

Initial Acceptance: 12/04/14
Final Acceptance: 01/12/15