Validity of the Student Risk Screening Scale: Evidence of Predictive Validity in a Diverse, Suburban Elementary Setting

Holly M. Menzies and Kathleen Lynne Lane

Journal of Emotional and Behavioral Disorders 2012 20: 82 originally published online 4 November 2011
DOI: 10.1177/1063426610389613

Hammill Institute on Disabilities

Journal of Emotional and Behavioral Disorders
20(2) 82–91

DOI: 10.1177/1063426610389613
http://jebd.sagepub.com
SAGE

Validity of the Student Risk Screening Scale: Evidence of Predictive Validity in a Diverse, Suburban Elementary Setting

Holly M. Menzies[1] and Kathleen Lynne Lane[2]

Abstract

In this study the authors examined the psychometric properties of the *Student Risk Screening Scale* (SRSS), including predictive validity in terms of student outcomes in behavioral and academic domains. The school, a diverse, suburban school in Southern California, administered the SRSS at three time points as part of regular school practices during one academic year to (a) monitor the overall level of risk evident in the building and (b) identify students at risk for antisocial behavior who might benefit from secondary (Tier 2) supports. Results suggest strong internal consistency (.85–.87) and test–retest stability (.69–.79). Results of predictive validity analyses suggest that initial ratings of risk as measured by the SRSS were predictive of teacher ratings of students' self-control skills and overall proficiency in language arts, with a statistically significant, inverse relationship. Students with higher levels of risk at the onset of the academic year were likely to end the year with lower levels of self-controls skills and lower levels of proficiency in language arts skill. SRSS scores were also predictive of office discipline referral (ODR) rates, with higher risk associated with higher rates of ODRs at year-end. Limitations and future directions are offered.

Keywords

systematic screening, antisocial behavior, Student Risk Screening Scale, schoolwide positive behavior support

Currently administrators and teachers are under tremendous pressure to support and educate students who enter the school system with greatly varying academic, behavioral, and social skill sets (Walker, Ramsey, & Gresham, 2004). For example, some elementary-age students begin their educational careers with strong early literacy skills (e.g., letter–sound and word identification skills) and social skills (e.g., cooperation, self-control) that facilitate the instructional process and enhance their ability to learn. In contrast, other students come to school with limited early literacy skills and antisocial behavior patterns that pose significant challenges for administrators, teachers, and their peer groups (Kauffman & Brigham, 2009; Lane, Menzies, & Kalberg, 2010). The term *antisocial behavior* refers to the opposite of prosocial behavior, namely, behaviors that include consistent violations of social norms (e.g., aggression, coercion). Not surprisingly, antisocial behavior is a foundational characteristic of students with emotional and behavioral disorders (EBD; Stouthamer-Loeber & Loeber, 2002; Walker et al., 2004).

Conservative estimates suggest that between 3% and 6% of school-age children and youth demonstrate these behavior patterns, placing them at heightened risk for negative outcomes within (e.g., school failure, impaired social relationships) and beyond (e.g., criminality, ongoing mental health needs; Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005) the school setting. It is important to note that not all students with antisocial behavior tendencies will require or receive special education services under the label of emotional disturbances (ED) as defined in the Individuals with Disabilities Education Improvement Act (2004). In fact, less than 1% of school-age students will be classified as having ED. As such, the general education community must be prepared to identify and support students with antisocial behavior to improve educational and postschool outcomes for this often difficult to teach population.

[1]California State University, Los Angeles, Los Angeles, CA, USA
[2]University of North Carolina, Chapel Hill, NC, USA

Corresponding Authors:
Holly M. Menzies, California State University, Los Angeles, Charter College of Education, 5151 State University Drive, Los Angeles, CA 90032
Email: hmenzies@exchange.calstatela.edu

Kathleen Lynne Lane, School of Education, University of North Carolina at Chapel Hill, 201-E Peabody Hall, Campus Box 3500, Chapel Hill, NC 27599
Email: Kathleen.lae@unc.edu

Identifying and Supporting All Student Using Three-Tiered Models of Prevention

To meet the multiple needs of all students, including those with and at risk for antisocial behavior, many schools across the country are utilizing a systems-level approach that includes progressively more intensive levels of support to (a) prevent harm from occurring (primary, schoolwide, or Tier 1 prevention), (b) reverse harm (secondary, selected, or Tier 2 prevention), and (c) reduce harm (tertiary, targeted, or Tier 3 prevention; Lane, 2007; Sugai & Horner, 2001). One critical component of such three-tiered models of prevention is accurate identification of which students require support beyond primary prevention. In other words, it is important that accurate, data-based decisions be made about who does and does not need secondary or tertiary levels of prevention (Kauffman, 2010). Every effort should be made to avoid either of the following errors: (a) identifying students who do not require additional supports (false positives) to eliminate unnecessary expenditure of resources and (b) overlooking students who will benefit from additional supports (false negatives), consequently missing an opportunity to prevent potentially deleterious outcomes.

Thus, an important component of three-tiered models of prevention is systematic screening of academic and behavioral performance. Although many schools regularly employ systematic screenings of academic performance such as *Dynamic Indicators of Basic Early Literacy Skills* (Kaminski & Good, 1996), analyzing benchmark scores and using data to formulate Tier 2 academic interventions, behavior screening tools are far less utilized (Lane, Menzies, & Kalberg, in press). Yet given the serious negative consequences of antisocial behavior and the fact that students are more amenable to intervention efforts during earlier stages of development (Bullis & Walker, 1994; Lane, 2003), it is essential that administrators, general education teachers, and parents actively engage in systematic screening and early intervention efforts to identify and support students exhibiting early signs of antisocial behavior. Such screening efforts also allow supports to be delivered in the most economical manner, before behavior patterns become firmly engrained and more resistant to intervention efforts (Kauffman & Landrum, 2009; Nelson, Babyak, Gonzalez, & Benner, 2003). Furthermore, data from behavior and academic screening tools can be analyzed in tandem to better address students' multiple needs with the three-tiered models of prevention described previously.

Core Characteristics of Systematic Screening Tools

Fortunately, there are a number of screening tools available for use at the elementary level such as the *Systematic Screening for Behavior Disorders* (SSBD; Walker & Severson, 1992; description to follow), the *Student Risk Screening Scale* (SRSS; Drummond, 1994; description to follow), the *Strengths and Difficulties Questionnaire* (SDQ; Goodman, 1997), the Social Skills Improvement System-Performance Screening Guide (SSiS-PSG; Elliot & Gresham, 2007), and the BASC2 Behavior and Emotional Screening System (BASC2 BESS; Kamphaus & Reynolds, 2007) see Lane, Menzies, Oakes, & Kalberg, 2010, for detailed descriptions of each measure).

When making decisions about which tool to adopt as part of regular school practices, it is important to select an instrument that has strong psychometric properties and is also feasible with respect to preparation, administration, scoring, and interpretation (Lane, Kalberg, Lambert, Crnobori, & Bruhn, 2010). The first goal is be certain that the tool is accurate—that it correctly identifies those students who do and do not have the condition of interest (in this case antisocial behavior). In brief, the goal is to limit false positive and false negatives in addition to establishing valid and reliable cut scores. The validity of a screening tool focuses on the evidence (empirical and logical) that supports the recommended use (and interpretation) of test scores (American Educational Research Association, 1999). Yet before an instrument can be considered valid, the reliability must be established. Reliability refers to the degree to which repeated administrations of the tool yield the same or similar results (Hatcher & Stepanski, 1994). Core characteristics of psychometrically sound systematic screeners include (a) high internal consistency ($\alpha \geq .80$), (b) high test–retest stability between scores obtained in consecutive administrations of the same instrument, (c) convergent validity with established measures of the same construct conducted at the same time, (d) positive predictive power (the probability that a student above the cut score is part of the target group with the construct; e.g., antisocial behavior) and negative predictive power (the probability that a person below the cut score is part of the group without the construct; e.g., without antisocial behavior), and (e) specificity (proportion of the reference group not identified given the same cut score) and sensitivity (proportion of the target population correctly identified; Kraemer, 1992; Lane, Kalberg, et al., 2010; Lane, Parks, Kalberg, & Carter, 2007; Lanyon, 2006).

The second goal is to select a screener that is feasible. No matter how strong a screener is psychometrically, it is far less likely to be administered accurately (or at all) if it is not feasible. A tool that is too lengthy, time-consuming to prepare or score, or difficult to interpret is unlikely to be adopted as part of regular school practices (Lane, Bruhn, Eisner, & Kalberg, 2010). Although teachers are highly committed to support students, they are confronted with a large number of multiple task demands each day, making time and energy precious commodities. Thus, the most

likely screener is one that is both reliable and valid yet also feasible.

Evaluating Screening Tools

Consider these dimensions as we examine two such tools: the SSBD and the SRSS. The SSBD is considered the gold standard for tools used in systematic screening to identify students with externalizing (e.g., aggression, coercion) and internalizing (e.g., shy, anxious, and withdrawn) behaviors. It includes a three-stage multiple gating system that begins with teacher rankings and ratings. It is followed by more intensive assessments such as rating scales for the six students (three with externalizing and three with internalizing behaviors) who pass through the first gate to Stage 2 (rating scales: *Critical Events Index* [CEI] and *Combined Frequency Index* [CFI]) and direct observations for the even fewer students who pass through the second gate to Stage 3 (direct observations). The instrument is inexpensive, costing less than $200 to purchase (including reproducible materials). In terms of administration, the first two stages take less than 1 hr to complete, with additional time required to score the CEI and CFI during Stage 2. One limitation of the SSBD it that scoring is somewhat laborious, and it does not account for students who share features of both externalizing *and* internalizing behavior patterns (comorbidity; Lane, Kalberg, et al., 2010). Given that these behaviors often tend to co-occur (e.g., students might have characteristics behavior patterns that include aggression and depression; Achenbach, 1991), this latter concern is a limiting feature of the SSBD.

The SRSS is a no-cost mass screening tool developed to detect students with antisocial behavior patterns. As part of this one-stage, one-page screener, teachers rate each student in their class on seven items using a 4-point Likert-type scale. Administration time is brief, requiring approximately 10 min for teachers to rate an entire class. Item-level data are summed into a total score, which is used to determine individual levels of risk as established by the test developer: 0–3 = low risk, 4–8 = moderate risk, and 9–21 = high risk (Drummond, 1994). Each student's individual score can be monitored over time to identify shift in risk status (Lane, Kalberg, Bruhn, Mahoney, & Driscoll, 2008). Although recent validity studies suggest that the SRSS does improve chance predictions of internalizing behaviors by as much as 30% (Lane, Little, et al., 2009), it is important to note that the SRSS was not designed to detect students with internalizing behaviors (e.g., depression, anxiety). The SRSS was constructed to detect students with antisocial behavior, with items reflective of externalizing behaviors and, to a lesser extent, covert behaviors (e.g., stealing).

As is evident from the descriptions, these measures vary in respect to the time required for preparation, administration, scoring, and interpretation. Collectively, these factors influence the social validity of each instrument, with the SRSS perhaps being more feasible than the SSBD in light of the greater ease of these logical considerations. In addition, the psychometric properties of each tool must be considered.

Lane, Little, et al. (2009) examined the degree to which the SRSS screening tool is equally sensitive and specific in identifying students with externalizing and internalizing behaviors as measured by the SSBD. This study of 562 kindergarten through second-grade students examined the concurrent validity of the SRSS to predict results from the SSBD when used to detect school children with externalizing or internalizing behavior concerns. When comparing students' SRSS scores in the low versus high risk categories, the SRSS was highly accurate for predicting both externalizing (95%) and internalizing (93%) problems on the SSBD. Sensitivity (94%) and specificity (95%) were both excellent for externalizing behavior. As expected, for internalizing behavior, sensitivity was lower (44%) whereas specificity was excellent (95%). Analysis of receiver operating characteristics curves also suggests that the SRSS is more accurate for detecting externalizing (area under the curve [AUC] = .952) than internalizing behaviors (AUC = .802). Collectively, findings provided initial evidence to support the use of the SRSS as an equally reliable tool as the SSBD for detecting students at risk for externalizing behavior. The SRSS does not share the same predictive accuracy as the SSBD for students with internalizing characteristics, as chance estimates improved by only 30%. Yet results must be interpreted in light of sample size limitations (e.g., only K–2 students, not all teachers participated, and not all parents allowed their children to participate).

To address these limitations, Lane, Kalberg, and colleagues (2010) replicated the Lane, Little, et al. (2009) study to further compare the SSBD and the SRSS. In this study of 2,588 students in middle Tennessee, the authors examined the psychometric properties of the Student SRSS. This included an evaluation of the concurrent validity of the SRSS to predict results from the SSBD when used to detect elementary students with externalizing or internalizing behavior concerns at three assessment points during one academic year. In this study, participating schools completed the SRSS and SSBD as part of regular school practices; therefore, all students in all grade levels were screened. Results suggest strong internal consistency (.81, .82, and .81) at the fall, winter, and spring administrations. Furthermore, test–retest stability analyses indicate that the SRSS is a stable instrument, with Pearson correlation coefficients ranging from .68 to .74. Analyses of receiver operating characteristics curves again suggested that the SRSS is more accurate for detecting externalizing than internalizing behaviors as measured through Stage 2 of the SSBD. Results demonstrate that the prediction of students with externalizing behaviors exceeds chances estimates by

45% to 46% (AUC values of .95, .95, and .96 at fall, winter, and spring administrations), whereas the prediction of students with internalizing behaviors exceeds chance estimates by 26%, 28%, and 32% across the three administrations of the SRSS.

Although these studies provide strong support for using either the SSBD or SRSS tools to detect externalizing behaviors, none of the validity studies conducted within the past decade have explored the extent to which these measures predict *academic* as well as behavioral outcomes for students with externalizing or antisocial behaviors. This article is a first step in exploring the current predictive validity of the SRSS to determine its accuracy in predicting behavioral *and* academic outcomes for elementary students.

Reliability studies of the SRSS conducted at the middle and high school levels suggest that SRSS scores predict grade point average (GPA) in middle and high schools as well as course failures at the middle school level (Lane, Bruhn, et al., 2010; Lane et al., 2007; Lane, Kalberg, Parks, & Carter, 2008). For example, in the Lane et al. (2007) study of 500 middle school students (Grades 8–9) living in a rural community, academic variables (i.e., GPA, number of course failures) could discriminate between students with moderate or high risks and students with low risk. Similar findings of predictive validity were also noted in a sample of urban middle school students. Students at low risk could be differentiated from those with moderate to high risk status on behavioral and academic measures (Lane, Bruhn, et al., 2010). Two years following the initial SRSS status, students in the low-risk group had significantly fewer out-of-school suspensions, fewer unexcused absences, and higher GPAs than students in the moderate- and high-risk groups. However, no such studies have been conducted in the past decade to determine if this measure predicts academic performance of academic outcomes at the elementary level.

Purpose

In this article, we report findings of a study conducted to extend the current literature on the reliability and validity of the SRSS for use in elementary schools. First, we examined the internal consistency of the SRSS in a diverse, suburban elementary school in Southern California at each of three administrations (fall, winter, and spring) during one academic year. Second, we established test–retest stability between these three administrations. Third, we used regression analyses to establish the predictive validity of the SRSS scores. Specifically, we conducted regression analyses to determine how well initial SRSS scores as rated by teachers at the onset of the school year predicted students' year-end (a) social competence as measured by teacher ratings of self-control using the *Social Skills Rating System* (Gresham & Elliott, 1990), (b) behavior performance as measured by office discipline referrals (ODRs), (c) attendance patterns as measured by the rate of absenteeism, and (d) academic performance as measured by district reading prompts.

Table 1. Participant Characteristics

		Sample (N = 286)	
Variable	Level	Frequency	Percentage
Gender	Boys	149	52.10
	Girls	137	47.90
Grade level	Kindergarten	55	19.23
	First	45	15.73
	Second	49	17.13
	Third	40	13.99
	Fourth	38	13.29
	Fifth	30	10.49
	Sixth	29	10.14

There were no missing data for gender or grade level.

Method

Participants

Participants were 286 students (149 [52.10%] boys, 137 [47.90%] girls) attending a diverse, suburban elementary school in Southern California who were rated by their teachers (N = 15) on the SRSS (description to follow) as part of regular school practices. The school served students in kindergarten through sixth grade (see Table 1). According to the state report card, the student body was ethnically and culturally diverse, with 33% Caucasian, 53% Hispanic, 12% African American, 1% Pacific Islander, and 1% Other. Of the students, 78% received free or reduced-price lunches. A chi-square analysis contrasting grade level × gender was not significant, $\chi^2(6, N = 286) = 12.51$, $p = .0515$.

Procedures

After participating in a professional development training series on designing, implementing, and evaluating an integrated three-tier model of prevention, the faculty elected to implement one systematic screening tool, the SRSS, as part of regular school practices. After obtaining university and district approvals, deidentified data were collected at the end of the academic year to (a) conduct an evaluation of the school's integrated model of prevention and (b) examine the psychometric properties of the SRSS (presented in the current article). As part of their assessment plan, the school-site leadership team administered the SRSS three times

during the academic year: 6 weeks into the school year (fall, Time 1), prior to winter break (winter, Time 2), and 6 weeks prior to year-end (spring, Time 3). These data were used to monitor the overall level of risk present in the building. Furthermore, these data were analyzed in conjunction with academic screening data (a district-developed curriculum-based measurement system) to identify students for secondary and tertiary levels of prevention (see Lane & Menzies, 2003, 2005).

The SRSS screening tools were introduced to the full faculty during a regularly scheduled faculty meeting by the reading specialist. She distributed the one-page screening instrument, provided directions on how to complete the measure, and explained how the data would be used to provide secondary and tertiary supports. Classroom teachers completed the measure in approximately 10 min, with the reading specialist and other members of the school-site team available to answer any questions. Completed screeners were turned in at the end of each of the three faculty meetings during which screening tools were completed. If a teacher was absent, he or she was asked to complete the screening tool within 1 week of the faculty meeting. Data were entered into an Excel database by the reading specialist, who also checked reliability of entry of all screenings. Any data entry errors (< 2%) were reconciled prior to data analysis. In this article, we analyzed the data for the 286 students enrolled during the academic year. Although there were no missing data for demographic data, there were missing data for the remaining measures. Analyses were conducted based on the data available.

Measures

Measures in the article include the SRSS, teacher ratings of social competence (specifically self-control skills), ODRs, rates of absenteeism, and language arts proficiency levels.

SRSS. The SRSS is a seven-item universal screening tool designed to identify elementary students at risk for antisocial behaviors patterns. Teachers rate each student in their class on 7 items—steal; lie; cheat; sneak; behavior problems; peer rejection; low academic achievement; negative attitude; and aggressive behavior—on a 4-point Likert-type scale ranging from 0 to 3 (*never* = 0, *occasionally* = 1, *sometimes* = 2, *frequently* = 3). A total score is computed for each student by summing the items, with total scores ranging from 0 to 21. Higher scores indicated higher levels of antisocial behavior. Drummond (1994) established three validated risk categories based on students' total scores: low (0–3), moderate (4–8), and high risk (9 or more).

The SRSS is a no-cost mass screening tool initially developed for use at the elementary level to distinguish between students who do and do not demonstrate antisocial tendencies (Drummond, Eddy, & Reid, 1998a, 1998b). Validity studies conducted by Drummond (the author) indicated that SRSS scores (a) correlated positively (r = .79) with the Aggressive Behavior subscale of the *Child Behavior Checklist* (Achenbach, 1991) and (b) predicted negative academic and behavioral outcomes from 1.5 to 10 years later (Drummond, Eddy, Reid, & Bank, 1994).

Self-control ratings. When designing the primary prevention plan, teachers rated students using the 30 items constituting the Social Skills subscale on the *Social Skills Rating System* (Gresham & Elliott, 1990). These 30 items are distributed equally across three factor analytically derived domains: cooperation, assertion, and self control. Teachers rated each item on a 3-point Likert-type scale: *not important* (0), *somewhat important* (1), *critical for success* (2). As part of the evaluation study, these data were analyzed to determine which domain to focus on as part of primary prevention efforts. Items constituting the self-control domain were identified as the most essential for success relative to the other two domains. Consequently, this domain became a core component of primary prevention efforts and was also identified as a performance measure.

To measure students' performance in self-control skills, teachers evaluated each student in their homeroom class on the 10 items constituting the self-control domain at year-end (e.g., "controls temper in conflict situation with adults," "receives criticism well"). Teachers rated students on a 3-point, Likert-type scale: *never* (0), *occasionally* (1), *sometimes* (2). Total scores were computed by summing the individual items for a total score that ranged from 0 to 20, with higher scores indicating higher levels of self-control skills according to teacher perceptions. Internal consistency estimates on the current sample for this measure are .92.

ODRs. ODRs were determined by computing rate. Specifically, the number of ODRs earned during the given academic year was divided by the number of days each student was enrolled to determine the rate of ODRs.

Absenteeism. Absenteeism was determined by computing a rate, specifically, the number of days the student was absent from school during the given academic year divided by the number of days the student was enrolled.

Language arts proficiency scores. Literacy skills were measured using three different measures, one for kindergarten, one for primary (1–2), and one for upper elementary (3–6) students. Kindergarten students' skills were assessed by measuring letter–sound proficiency using a district-developed multiple measure assessment.

Primary students' skills were assessed using *Scholastic Comprehension Tests* (Scholastic, 1996). This measures comprehension, study skills, and grammar, usage, and mechanics. The reading domain includes 20 short-answer

Table 2. Cronbach's Coefficient Alpha—Total and With Deleted Variables

Time	Item	*M*	*SD*	Standardized variables: Correlation with total	Standardized variables: α
Fall (*n* = 252)					.87
	1. Steal	0.10	0.39	.38	.89
	2. Lie, cheat, sneak	0.75	1.02	.79	.83
	3. Behavior problem	0.91	1.07	.77	.84
	4. Peer rejection	0.43	0.86	.59	.86
	5. Low academic achievement	1.14	1.09	.52	.87
	6. Negative attitude	0.63	1.00	.74	.84
	7. Aggressive behavior	0.52	0.93	.77	.84
Winter (*n* = 271)					.85
	1. Steal	0.10	0.35	.43	.86
	2. Lie, cheat, sneak	0.66	0.94	.70	.82
	3. Behavior problem	0.83	1.03	.73	.81
	4. Peer rejection	0.31	0.76	.58	.84
	5. Low academic achievement	1.15	1.21	.50	.85
	6. Negative attitude	0.55	0.95	.69	.82
	7. Aggressive behavior	0.42	0.85	.68	.82
Spring (*n* = 269)					.86
	1. Steal	0.22	0.59	.48	.86
	2. Lie, cheat, sneak	0.74	1.02	.79	.82
	3. Behavior problem	0.88	1.07	.74	.83
	4. Peer rejection	0.31	0.69	.62	.85
	5. Low academic achievement	1.13	1.15	.39	.88
	6. Negative attitude	0.55	0.96	.71	.83
	7. Aggressive behavior	0.52	0.92	.72	.83

and multiple-choice items measuring skills and strategies in comprehension, vocabulary, and literacy elements. An additional 10 items measure grammar, usage, and mechanics skills, and another 5 items assess study skills.

Upper elementary students' literacy skills were assessed using *Harcourt Brace Reading Comprehension Tests* (Harcourt Brace, 1997). These criterion-referenced grade-level tests contain two reading passages, each of which contains two short-answer and eight multiple-choice questions. Selections are drawn from grade-level literature and are designed to determine the extent to which a student's reading comprehension skills are at grade level.

For kindergarten, lower, and upper elementary students, district-level officials convert raw scores to proficiency levels as follows: *outstanding* = 4, *proficient* = 3, *basic* = 2, and *below basic* = 1.

Statistical Analysis

Internal consistency of the SRSS was examined by computing alpha coefficients for each administration (fall, winter, and spring). Test–retest stability was examined by computing Pearson correlation coefficients between the (a) fall and winter, (b) fall and spring, and (c) winter and spring time points. Regression analyses were conducted to establish the predictive validity of the SRSS scores. Specifically, we conducted regression analyses to determine how well initial SRSS scores (a continuous variable; $X_{\text{SRSS-total score}}$) as rated by teachers at the onset of the school year predicted students' year-end (a) social competence as measured by teacher ratings of self-control using the *Social Skills Rating System* ($Y_{\text{self-control}}$), (b) behavior performance as measured by ODRs (Y_{ODRs}), (c) attendance patterns as measured by the rate of absenteeism (Y_{absences}), and (d) academic performance as measured by district reading prompts (Y_{reading}). In addition, we computed correlation coefficients (r) to evaluate the strength of the relationship between SRSS total scores ($X_{\text{SRSS-total score}}$) and each predictor variable ($Y_{\text{self-control}}$, Y_{ODRs}, Y_{absences}, and Y_{reading}).

Results

Internal Consistency

Based on teach ratings of the entire study body at three points, (fall, winter, and spring), Cronbach's alpha values were greater than or equal to .85, with respective values of .87, .85, and .86 (see Table 2; see Table 3 for intercorrelations).

Table 3. Intercorrelations

Time	Item	1	2	3	4	5	6	7
Fall (*n* = 252)	1. Steal	1.00						
	2. Lie, cheat, sneak	.39	1.00					
	3. Behavior problem	.35	.77	1.00				
	4. Peer rejection	.20[a]	.50	.46	1.00			
	5. Low academic achievement	.19[b]	.46	.49	.43	1.00		
	6. Negative attitude	.35	.65	.67	.54	.48	1.00	
	7. Aggressive behavior	.36	.74	.71	.57	.40	.64	1.00
Winter (*n* = 271)	1. Steal	1.00						
	2. Lie, cheat, sneak	.33	1.00					
	3. Behavior problem	.30	.68	1.00				
	4. Peer rejection	.51	.45	.43	1.00			
	5. Low academic achievement	.23[c]	.43	0.43	.33	1.00		
	6. Negative attitude	.27	.60	.63	.39	.48	1.00	
	7. Aggressive behavior	.33	.56	.67	.45	.34	.62	1.00
Spring (*n* = 269)	1. Steal	1.00						
	2. Lie, cheat, sneak	.58	1.00					
	3. Behavior problem	.35	.73	1.00				
	4. Peer rejection	.34	.53	.50	1.00			
	5. Low academic achievement	.18	.33	.39	.30	1.00		
	6. Negative attitude	.40	.65	.66	.51	.29	1.00	
	7. Aggressive behavior	.35	.63	.64	.60	.33	.64	1.00

All correlations are significant at the $p < .0001$ level, except for a = .0014, b = .0031, c = .0001.

Test–Retest Stability

Test–retest stability was computed for ratings provided at fall and winter (approximately 15 weeks in duration), fall and spring (approximately 43 weeks in duration), winter and spring (approximately 28 weeks in duration). Results revealed a statistically significant correlation between total scores as follows: fall to winter, $r = .78$, $p < .0001$; fall to spring, $r = .69$, $p < .0001$; winter to spring, $r = .79$, $p < .0001$ (see Table 4).

Table 4. Test–Retest Stability

Time frame comparison time	*n*	Weeks	Correlation coefficient
Fall–winter	225	15	.78
Fall–spring	202	43	.69
Winter–spring	220	28	.79

All correlation coefficients are significant at the $p < .0001$ level.

Predictive Validity

Regression analyses indicate that fall SRSS total scores predicted year-end performance for three of the four variables: self-control, $F(1, 187) = 98.68$, $p < .0001$, accounting for 34% of the variance; ODRs, $F(1, 200) = 61.10$, $p < .0001$, accounting for 23% of the variance; and reading performance, $F(1, 198) = 10.71$, $p = .0011$, accounting for 5% of the variance (see Table 5). Correlations coefficients suggest a strong inverse relation between initial SRSS scores and year-end teacher ratings of self-control ($r = -.59$, $p < .0001$) as well as a low to moderate significant relation with year-end reading performance ($r = -.23$, $p = .0013$). This suggest that students with higher levels of risk as measured by the SRSS were likely to have lower levels of self-control skills and lower proficiency in language arts skills at year-end. Furthermore, there was a significant relation between SRSS scores and year-end ODR ($r = .48$, $p < .0001$), indicating that students with higher levels of risk at the onset of the year were more apt to earn ODRs relative to students who had initially lower levels of risk. There was not a significant relation between SRSS ratings and attendance patterns.

Discussion

Systematic screenings for academic and behavior performance provide administrators and teachers with reliable methods for determining which students do—and do not—require additional support (e.g., secondary and tertiary levels of prevention) within the context of three-tier models of prevention (Lane, Menzies, & Kalberg, 2010). Although a number of systematic screening tools are available for use at the elementary level, predictive validity studies have focused exclusively on predicting important behavioral outcomes

Table 5. Predictive Validity of Student Risk Screening Scale (SRSS) Scores: Behavioral and Academic Outcomes

Domain	Variable	Correlation with Time 1 SRSS total score (significance)	*F* value	Significance testing	R^2
Behavioral	Self-control ratings	−.59 ($p < .0001$)	$F(1, 187) = 98.68$	$p < .0001$	.34
	Office discipline referrals	.48 ($p < .0001$)	$F(1, 200) = 61.10$	$p < .0001$	.23
	Absenteeism	.04 ($p = .5698$)	$F(1, 198) = 0.32$	$p = .57$	.00
Academic	Language arts proficiency	−.23 ($p = .0013$)	$F(1, 187) = 10.71$	$p = .0011$	.05

R^2 refers to the percentage of variance accounted for in the model.

(e.g., ODRs), with no attention to academic performance. To date, the only validity studies conducted in the past decade that have explored behavioral and academic outcomes are those that examined the utility of the SRSS at the middle (Lane, Bruhn, et al., 2010; Lane et al., 2007) and high school (Lane, Kalberg, Parks, et al., 2008) levels. The intent of this study was to extend this body of literature to determine if an early indication of antisocial behavior as measured by the SRSS could predict important behavioral, social, and academic outcomes during the elementary years. Moreover, we examined the psychometric properties of the SRSS, which was implemented as part of regular school practices in a diverse suburban school in Southern California, including predictive validity with respect to students' behavioral and academic outcomes.

In terms of psychometric properties, the internal consistency estimates of the SRSS ranged from .85 to .87 across the fall, winter, and spring administrations within the given academic year. These internal consistency estimates slightly exceed alpha coefficient reported in validity studies conducted in elementary schools located in the southern region of the United States (Lane, Kalberg, et al., 2010; Lane, Little, et al., 2009). Furthermore, estimates were consistent with alpha coefficients reports in middle and high schools in urban and rural settings in the southern part of the United States (Lane, Bruhn, et al., 2010; Lane et al., 2007; Lane, Kalberg, Parks, et al., 2008).

Analyses of test–retest stability were also consistent with findings in elementary, middle, and high school settings. Namely, the SRSS is a stable instrument as evidenced by Pearson correlation coefficients ranging from .69 to .79, which are highly similar to findings at the middle school level (.68 to .74; Lane, Bruhn, et al., 2010). These findings are also consistent with earlier psychometric studies of the SRSS reported by the developer of the instrument (Drummond) at the elementary level as well as more recent studies in elementary schools (Lane, Kalberg, et al., 2010; Lane, Little, et al., 2009).

Finally, results of predictive validity analyses suggest that initial ratings of risk as measured by the SRSS were predictive of the ODR rates, with higher risk associated with higher rates of ODRs at year-end. Again, these findings are consistent with validation studies conducted at the middle and high school levels. In addition, fall SRSS scores were predictive of year-end teacher ratings of students' self-control skills and overall proficiency in language arts, with a statistically significant, inverse relationship. Namely, students with higher levels of risk at the onset of the academic year were likely to end the year with lower levels of self-controls skills and lower levels of proficiency in language arts skills.

Collectively, these findings provide compelling—yet preliminary—evidence to suggest that this instrument is predictive of salient behavioral (ODR), social (teacher ratings of self-control skills), and academic (proficient in language arts skills) outcomes. Namely, teacher ratings as measured by the SRSS just 6 weeks after elementary students begin the school year can identify students who are apt to struggle in a number of important domains. It raises the question, if we can find these students in just 10 min at the beginning of the year, why wait for students to fail? Why wait for them to earn ODRs or perform below benchmarks in reading—a keystone skill that has the potential to unlock all learning? Effective screeners can be used to identify accurately students who will benefit from assistance in the form of secondary (Tier 2) and tertiary (Tier 3) supports provided in advance rather than offering remediation after a student has experienced failure.

As former teachers, we are cautious about asking teachers to complete any unnecessary task that adds to the myriad demands that they face on a daily basis. Yet given the predictive validity of the SRSS, we contend that completing such a measure is worth the investment.

If information gleaned from the SRSS can provide school-site teams and research teams with the necessary information to identify students who will struggle behaviorally, socially, and academically, we strongly recommend that such screening tools be adopted as part of regular school practices. Given the predictive accuracy and relatively little investment with respect to time and other resources, the SRSS is an effective tool for identifying students for secondary and tertiary supports relative to less reliable techniques (e.g., ODRS). Yet we also temper our enthusiasm with the recognition that this is but one study, which significantly limits the generalizability of the findings. Consequently, we recommend that readers consider the following limitations when interpreting outcomes.

Limitations and Future Directions

There are several key considerations that warrant attention when interpreting findings. First, the sample involves students and teachers from just one school in Southern California. It is mandatory that additional studies be conducted with a greater breadth of teacher and student participants to establish the generalizability of the results. Namely, data from elementary schools in other geographic regions and other locales must be collected.

Second, this study focuses solely on the SRSS. We recommend other research teams explore additional systematic screening tools, including the SSBD, the SDQ, the SSiS-PSG, and the BASC2 BESS, to determine if these instruments hold equal (or greater) predictive accuracy in behavioral, social, and academic domains. This additional information will be very useful to practitioners, particularly school-site leadership teams, as they make decisions regarding which behavior screening tools to adopt as part of regular school practices.

As mentioned at the beginning of this study, information on psychometric rigor and feasibility is critical if school-site teams are to make adoptions decisions. We contend that instruments lending clear direction as to which students may require secondary (Tier 2) and tertiary (Tier 3) levels of prevention within the context of three-tiered models of prevention will be highly useful (Lane, Kalberg, & Menzies, 2009).

Third, although the outcomes of this study are intriguing, the predictive analyses focused on outcomes within a single academic year. It is imperative that additional studies be conducted to assess the long-term predictive validity of the SRSS (e.g., 2, 5, and even 10 years following the initial nomination).

Fourth, as with the Lane, Bruhn, and colleagues (2010) study, we did not formally assess perceptions of the feasibility of the SRSS. We strongly recommend that future studies examine information regarding the effort (and acceptability of the effort) involved in preparing, administering, scoring, and interpreting screening tools. We believe that systematic screenings of academic and behavioral performance are essential to (a) monitor the overall level of risk present in a school and (b) identify students who may require secondary and tertiary supports (Lane, Kalberg, et al., 2009; Lane, Menzies, Oakes, & Kalberg, in press). To this end, we encourage the research and teaching communities to work collaboratively to identify screening tools that meet rigorous psychometric standards as well as adhere to reasonable expectations for what they can deliver.

Summary

Despite these limitations, we view this study as providing important initial evidence to support the predictive validity of the SRSS in predicting important behavioral, social, *and* academic outcomes for students. This instrument may hold utility for the identification of students who require additional supports within comprehensive three-tiered models of prevention. Because we hold the SSBD in the highest regard, we encourage other research teams to (a) replicate the findings presented in this study and (b) explore the predictive validity of the SSBD as well as other more recent screening tools (e.g., SSiS-PSG, BASC2 BESS) in predicting multiple outcomes for elementary-age students, particularly instructional outcomes given that instruction is a primary objective of all schools.

Declaration of Conflicting Interests

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

Funding

The authors received no financial support for the research and/or authorship of this article.

References

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

American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Bullis, M., & Walker, H. M. (1994). *Comprehensive school-based systems for troubled youth*. Eugene: University of Oregon, Center on Human Development.

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. (1998a). *Follow-up study #3: Risk Screening Scale: Prediction of negative outcomes by 10th grade from 2nd grade screening* (Unpublished technical report). Eugene: Oregon Social Learning Center.

Drummond, T., Eddy, J. M., & Reid, J. B. (1998b). *Follow-up study #4: Risk Screening Scale: Prediction of negative outcomes in two longitudinal samples* (Unpublished technical report). Eugene: Oregon Social Learning Center.

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

Elliott, S. N., & Gresham, F. M. (2007). *Classwide intervention program teacher's guide: Social Skills Improvement System*. Minneapolis, MN: NCS Pearson.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586.

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

Harcourt Brace. (1997). *Harcourt Brace Reading Comprehension Tests. Signatures*. New York, NY: Author.

Hatcher, L., & Stepanski, E. J. (1994). *A step-by-step approach to using the SAS system for univariate and multivariate statistics*. Cary, NC: SAS Institute.

Individuals with Disabilities Education Improvement Act, Pub. L. No. 108-446, 118 Stat. 2647 (2004).

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review*, *25*, 215–227.

Kamphaus, R. W., & Reynolds, C. R. (2007). *BASC-2 Behavior and Emotional Screening System (BASC-2 BESS)*. San Antonio, TX: Pearson.

Kauffman, J. M. (2010). *The tragicomedy of public education: Laughing and crying*, *thinking and fixing*. Verona, WI: Full Court.

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

Kauffman, J. M., & Landrum, T. (2009). *Characteristics of emotional and behavioral disorders of children and youth* (8th ed.). Columbus, OH: Merrill.

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Newbury Park, CA: Sage.

Lane, K. L. (2003). Identifying young students at risk for antisocial behavior: The utility of "teachers as tests." *Behavioral Disorders*, *28*, 360–389.

Lane, K. L. (2007). Identifying and supporting students at risk for emotional and behavioral disorders within multi-level models: Data driven approaches to conducting secondary interventions with an academic emphasis. *Education and Treatment of Children*, *30*, 135–164.

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.

Lane, K. L., Kalberg, J. R., Bruhn, A. L., Mahoney, M. E., & Driscoll, S. A. (2008). Primary prevention programs at the elementary level: Issues of treatment integrity, systematic screening, and reinforcement. *Education and Treatment of Children*, *31*, 465–494.

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.

Lane, K. L., Kalberg, J. R., & Menzies, H. M. (2009). *Developing schoolwide programs to prevent and manage problem behaviors: A step-by-step approach*. New York, NY: Guilford.

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.

Lane, K. L., Little, A. M., 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. (2003). A school-wide intervention with primary and secondary levels of support for elementary students: Outcomes and considerations. *Education and Treatment of Children*, *26*, 431–451.

Lane, K. L., & Menzies, H. (2005). Teacher-identified students with and without academic and behavioral concerns: Characteristics and responsiveness to a school-wide intervention. *Behavior Disorders*, *31*, 65–83.

Lane, K. L., Menzies, H. M., & Kalberg, J. R. (2010). An integrated, comprehensive three-tier model to meet students' academic, behavioral, and social needs. In K. Harris, T. Urdan, & S. Graham (Eds.), *American Psychological Association educational psychology handbook*. Washington, DC: American Psychological Association.

Lane, K. L., Menzies, H. M., Oakes, W. P., & Kalberg, J. R. (2010). *Systematic screenings of behavior to support instruction: From preschool to high school*. Manuscript under contract.

Lane, K. L., Oakes, W. P., & Menzies, H. M. (2010). Systematic screenings to prevent the development of learning and behavior problems: Consideration for practitioners, researchers, and policy makers. *Journal of Disabilities Policy Studies, 21*, 160–172.

Lane, K. L., Parks, R. J., Kalberg, J. R., & Carter, E. W. (2007). Student Risk Screening Scale: Initial evidence for score reliability and validity at the middle school level. *Journal of Emotional and Behavioral Disorders*, *15*, 209–222.

Lanyon, R. (2006). Mental health screening: Utility of the psychological screening inventory. *Psychological Services*, *3*, 170–180.

Nelson, N. J., Babyak, A., Gonzalez, J., & Benner, G. J. (2003). An investigation of the types of problem behaviors exhibited by K–12 students with emotional or behavioral disorders in public school settings. *Behavioral Disorders*, *28*, 348–359.

Scholastic. (1996). *Scholastic Comprehension Tests*. *Literacy place*. New York, NY: Author.

Stouthamer-Loeber, M., & Loeber, R. (2002). Lost opportunities for intervention: Undetected markers for the development of serious juvenile delinquency. *Criminal Behaviour and Mental Health*, *12*, 69–82.

Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy*, *24*, 25–50.

Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005). The children and youth we serve: A national picture of the characteristics of students with emotional disturbance receiving special education services. *Journal of Emotional and Behavioral Disorders*, *13*, 79–96.

Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices*. Belmont, CA: Wadsworth.

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