

2017

Measuring Teachers' Perceptions of Student Behavior Using the Systematic Screening for Behavior Disorders

Megan Ruth Erickson

Louisiana State University and Agricultural and Mechanical College, merickson76@gmail.com

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Psychology Commons](#)

Recommended Citation

Erickson, Megan Ruth, "Measuring Teachers' Perceptions of Student Behavior Using the Systematic Screening for Behavior Disorders" (2017). *LSU Master's Theses*. 4567.

https://digitalcommons.lsu.edu/gradschool_theses/4567

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

MEASURING TEACHERS' PERCEPTIONS OF STUDENT BEHAVIOR USING THE
SYSTEMATIC SCREENING FOR BEHAVIOR DISORDERS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

by

Megan Ruth Erickson
B.A., Rice University, 2012
May 2017

TABLE OF CONTENTS

ABSTRACT.....	iii
INTRODUCTION	1
METHOD	15
RESULTS	19
DISCUSSION.....	30
REFERENCES	42
APPENDIX.....	48
VITA.....	53

ABSTRACT

As part of school- and district-wide preventative efforts, universal screening serves to identify students at risk for emotional and behavioral disorders while their behavior is still amenable to treatment. However, there are few universal screeners available for middle school students, who may be at heightened risk for developing emotional and behavioral difficulties due to major academic and social changes. The *Systematic Screening for Behavior Disorders* (SSBD), often considered the “gold standard” for behavior screeners, was recently validated for use in middle schools. However, there is little research on the reliability and validity of the SSBD for screening middle school students and the differences in responses between teachers of different academic areas. The purpose of this study is to extend the previous research validating the SSBD by assessing the adequacy of its technical characteristics and sensitivity in middle school students among teachers of different academic subjects.

INTRODUCTION

Educating students with emotional and behavioral difficulties presents a significant challenge for teachers and school systems, resulting in disproportionately high rates of teacher stress, attrition, and loss of instructional time (Soodak, Podell, & Lehman, 1998; Cross & Billingsley, 1994). Emotional and behavioral disorders (EBD) are characterized by a pattern of maladaptive behaviors that significantly impede upon one's ability to function successfully in academic and interpersonal activities (Walker, Ramsey, & Gresham, 2004). The range of behaviors indicative of EBD are generally divided upon two dimensions: externalizing and internalizing behaviors (Kauffman, 2001). Externalizing behaviors are undercontrolled, often disruptive behaviors directed outwardly towards the environment (Kauffman, 2001; Achenbach & Edelbrock, 1978). In the school setting, examples of externalizing behaviors include defying the teacher's instructions, talking out of turn, and committing acts of violence and aggression.

Contrastingly, internalizing behaviors are overcontrolled behaviors directed inwardly towards the self, such as fear, sadness, or social withdrawal. (Kauffman, 2001; Achenbach & Edelbrock, 1978). Although the effects of externalizing behaviors on individuals and classrooms are more readily apparent, internalizing behaviors are no less serious: they can significantly interfere with a child's development of interpersonal relationships to the point where it hinders their acquisition of appropriate social skills (Walker et al., 2004). When patterns of internalizing and externalizing behavior become sufficiently detrimental to an individual's everyday function, they might meet criteria for related disorders in the *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition* (5th ed.; DSM 5; American Psychiatric Association, 2013) such as oppositional defiant disorder, conduct disorder, and major depressive disorder, or they could be classified as emotionally disturbed (ED) under the Individuals with Disabilities Education Improvement Act (American Psychiatric Association, 2013; IDEA, 2004).

Outcomes of Emotional and Behavioral Disorders

The characteristics of EBD have severe and pervasive effects on the children and youth who experience them. The relationship between problem behaviors and negative academic and social outcomes is well established in the literature (Malecki & Elliot, 2002). Research indicates that children with EBD exhibit low achievement in all academic subjects, perform approximately 1-2 years below grade level, and are more likely to be retained, suspended, or expelled than students of other special education categories (Nelson, Benner, Lane, & Smith, 2004; Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005; Trout, Nordess, Pierce, & Epstein, 2003; Wagner, 1995). Furthermore, these early academic and behavioral challenges have damaging effects on later academic outcomes: roughly half of students with EBD fail to finish high school and those who do are significantly less likely to attend postsecondary school (Wagner et al., 2005).

The substantial consequences of EBD continue to interfere with adaptive functioning post-high school. Long-term outcomes of ED classification were recorded in the National Longitudinal Transition Study-2, a U.S. Department of Education project that followed a representative sample of approximately 11,000 students with disabilities over 10 years (Wagner et al., 2005). As of eight years post-high school, 75% of those with ED had been involved with the criminal justice system in some regard. Specific data showed that 75% had been stopped by police for traffic or other minor violations, 60% had been arrested, and 37% had spent a night in jail. Additionally, only 50% of those with ED were employed, had a greater job turnover rate, and had an overall more difficult time adjusting to post-school adulthood than students with other disabilities (Wagner et al. 2005; Davis & Vander Stoep, 1997). Because of the pervasive and severe outcomes associated with EBD both within the individual and in larger academic and social systems, early identification and prevention efforts are key to these students' success.

Importance of Early and Accurate Identification

Fortunately, there are a number of evidence-based interventions for at-risk youth which successfully improve their behavioral deficits (Cook, Landrum, Tankersley, & Kauffman, 2003). However, of the approximately 20% of children exhibiting symptoms of EBD, only 1% are identified to receive school-based services under IDEA (Wagner et al. 2005). The low identification and treatment of at-risk students has in part to do with the inadequacy of traditional reactive referral methods (Walker, Nishioka, Zeller, Severson, & Feil, 2000; Kratochwill, Albers, & Shernoff, 2004). Commonly known as the “wait to fail” approach, students are referred for further assessment or intervention only after committing significant rule violations leading to office disciplinary referrals (ODRs), suspensions, or expulsion (Walker et al., 2000). This method represents a particularly serious problem for identifying students at risk for internalizing disorders. Internalizing behaviors such as depression, anxiety, social withdrawal, and somatic complaints are often non-disruptive and easily overlooked by teachers, rendering these students less likely to be referred based on traditional markers of risk (ODRs, etc.; Gresham & Kern, 2004; Lane et al. 2007; Kauffman, 1999). The shortcomings of reactive referral methods often result in a delay in the beginning of services (Kamphaus, DiStefano, Dowdy, Eklund, & Dunn, 2010). Numerous studies have found that behavior becomes increasingly resistant to change as a child ages, thus underlining the importance of developing alternative methods for identifying at-risk students (Loeber, 1991).

Proactive Methods and Universal Screening

The challenges posed by traditional referral methods highlight the need for early identification and prevention before behavior problems become a persistent element of a child’s classroom behavior. To this effect, some schools are implementing proactive models of positive

behavioral support (PBS) that emphasize school- and district-wide programs to promote a positive and effective learning environment (Sugai & Horner, 2006). Central to this and similar models (e.g. Response-to-Intervention [RTI]) is a system of multi-tiered support, in which intervention intensity increases based on the success or failure of students to respond to universal measures (Sugai & Horner, 2006). Primary prevention and intervention systems consist of school-wide efforts such as increased supervision, positive reinforcement systems, and posted school-wide behavioral expectations in which all students participate (Sugai & Horner, 2006). Approximately 15-20% of students continue to exhibit behavior problems despite primary interventions. These students are then selected for more targeted secondary interventions, which may include interventions such as small-group social skills instruction (Lane, Oakes, Menzies, & Germer, 2014). Finally, an estimated 1-5% of students require tertiary interventions, which are the most intensive and include highly individualized, function-based behavior plans designed and implemented by specialized professionals (Lane et al. 2014; Sugai & Horner, 2006).

Fundamental to multi-tiered systems is the early identification of students in need of secondary or tertiary interventions (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). To this end, an increasing number of schools are integrating universal screening into their PBS model, a process in which all students are rated on a short assessment for the purpose of identifying students at risk for internalizing and/or externalizing behavioral concerns (Albers & Kettler, 2014). The use of emotional and behavioral screeners is intended to combat the “wait-to-fail” approach by improving the rate at which students will be placed in more intensive interventions, while their behavior is still responsive to modification (Dowdy, Doane, Eklund, & Dever, 2011). Universal screening serves a “triage function,” by identifying which students will

benefit most from secondary or tertiary interventions (Nelson, Benner, Reid, Epstein, & Currin, 2002; Caldarella, Young, Richardson, Young, & Young, 2008).

Characteristics of Reliable and Valid Screeners

Currently, there are numerous universal screening measures available for assessing behavioral and emotional risks (Dowdy et. al, 2011; Lane, Kalberg, Lambert, Crnabori, & Bruhn, 2010). Because different measures may be more or less appropriate for different purposes and settings, several considerations should be made when choosing a screener (Lane et al., 2014). Glover and Albers (2007) outline important criteria to consider when selecting a screener, among which are the inclusion of norms for the target population, adequate reliability, and the validity with which a screener identifies at-risk individuals (Glover and Albers, 2007).

Norms are the collection of scores from a representative sample of the population to which an individual score can be compared (American Educational Research Association [AERA], American Psychological Association [APA], & National Council for Measurement in Education [NCME], 1999). To understand an individual score, it must be considered within the context of an appropriate reference group that approximates the population in terms of gender, age, race, socioeconomic status, and other relevant characteristics (Phillips, 1982; Taylor, 2000). When selecting a screening measure, test users should consider the norming data to ensure their accuracy in interpreting results (Glover & Albers, 2007).

Another important consideration when choosing a behavior screener is its reliability, traditionally defined as the consistency of scores across multiple administrations, raters, or items (AERA et al., 1999). Glover and Albers (2007) outline three main types of reliability: internal consistency, test-retest, and inter-rater reliability. Internal consistency coefficients represent the relationship between individual items within the same measure (AERA et al., 1999). These

coefficients, such as Cronbach's alpha, indicate whether various items measure the same construct, such as internalizing or externalizing behaviors (Lane et al., 2014). Test-retest reliability refers to the stability of scores across multiple administrations (AERA et al., 1999; Glover & Albers, 2007). High test-retest reliability indicates that a measure is less susceptible to error due to measurement or change over time (Glover & Albers, 2007). Finally, inter-rater reliability is the uniformity with which different scorers rate an individual's responses or behaviors (Lane et al., 2014; AERA et al., 1999). For a test to have adequate inter-rater reliability, there should be a strong correlation between scores from different raters. Because universal screening results in a low-stakes decision, recommended reliability coefficients should be between .70 and .80 on a scale from 0.0-1.0 (Albers & Kettler, 2014).

While reliability is a necessary component of any assessment, scores are meaningless without evidence-based validation for their interpretation and use (AERA et al., 1999). Validity is traditionally defined as the extent to which appropriate inferences can be made based on the results of an assessment, or the degree to which an assessment measures what it purports to measure (Messick, 1987; Glover & Albers, 2007). A critical aspect of a universal screener's validity is its sensitivity and specificity (Glover and Albers, 2007). Sensitivity (or true positives) refers to the correct identification of students who are at risk; conversely, specificity (or true negatives) refers to the correct identification of students who are *not* at risk (Lane et al., 2014). A low sensitivity index indicates that at-risk students are being overlooked and thus hazard the possibility of not obtaining needed services (Glover and Albers, 2007). Similarly, a low specificity index implies that students are being over-identified for services, which unnecessarily consumes resources for students who do not need secondary or tertiary interventions (Algozzine and Ysseldyke, 1986). Valid cutoff scores should minimize over- and under-identification with

recommended sensitivity and specificity values of 75-80% (Kingslake, 1983; Carter, Briggs-Gowen, & Davis, 2004).

An additional type of validity is concurrent validity, which represents the strength of a relationship between two tests that measure the same criterion (Glover and Albers, 2007). Concurrent validity can best be understood as the degree to which a measure can identify students who are currently experiencing behavior problems (Glover & Albers, 2007). Finally, construct validity indicates whether a test measures the theoretical variable, or construct, that it intends to measure (Glover and Albers (2007). Two important types of construct validity are convergent and discriminant validity, which are the degrees of correlation between measures that purportedly measure the same or different constructs respectively (AERA et al., 1999). Because it focuses on the relationship between new and established assessments of constructs, convergent validity is sometimes considered the most important subtype of validity evidence for universal screeners (Albers & Kettler, 2014). Because different types of validity measure distinct criteria for evaluating the effectiveness of an assessment, multiple validity measures should be taken into account when choosing a screener.

Adolescent Development and Screening in Middle Schools

While many reliable and valid emotional and behavior screening systems exist for elementary school students, there is a notable shortage of empirically validated measures for use in middle and junior high schools (Walker, Cheney, Stage, & Blum, 2005; Lane, Wheby, Robertston, & Rogers, 2007). Unfortunately, the transition to middle school is characterized by a number of unique and challenging social and academic changes, which can heighten the risk for developing EBD (Eccles, Lord, & Midgley, 1991). Peer interaction becomes more frequent and less likely to be supervised by adults (Brown, 1990; Rubin, Bukowski, & Parker, 1998) and

adolescents increasingly depend on their friends for acceptance rather than on their parents for approval (Steinberg & Silverberg, 1986). Larger schools, the onset of puberty, and budding romantic relationships contribute to declines in academic achievement, self-esteem, extracurricular participation, and academic engagement ((Eccles et al. 1991; Holas & Huston, 2012, Gutman & Midgley, 1999). Furthermore, the relationships between teachers and their students begin to suffer, as teachers tend to become more discipline-oriented and less trusting of their students, resulting in the loss of the close teacher-student connections that often serve as protective factors in increased externalizing and internalizing behaviors (Midgley, Feldlaufer, & Eccles, 1989; Holad & Huston, 2012; Alspaugh, 1998). Additionally, students must adjust to added high-stakes testing, reduced instructional time, and juggling the expectations of multiple teachers (Eccles et a., 1991; Midgley, Middleton, Kumar, & Gheen, 2002).

Any combination of these factors can increase the risk for EBD, particularly for emotionally vulnerable students (Eccles et al., 1991). Approximately 50% of children with EBD are estimated to develop the disorder after age 11 (Merikangas, et al., 2010), and the rate of ED identification peaks in grades 9 and 10, far beyond the developmental period ideal for behavioral intervention (Walker et al., 2000). While late-onset behavior disorders are typically associated with better outcomes than early-onset disorders, the significant impact of EBD on long-term academic and social functioning still underscores the need for reliable and valid screening measures for this population (Moffitt, 1993; Kauffman, 2001; Stouthamer-Loeber & Loeber, 2002). While there are currently available screening measures for early adolescents (e.g., SRSS), these measures may be too costly and time-consuming for schools to adopt (Caldarella et al., 2009). Fortunately, a highly reliable and valid measure has recently been approved for use with

early adolescents: the *Systematic Screening for Behavior Disorders, Second Edition* (SSBD; Walker, Sevenson, & Feil, 2014)

Systematic Screening for Behavior Disorders, Second Edition

The SSBD is a multiple-gated assessment procedure often considered the “gold standard” for behavior screeners (Lane, Menzies, Oakes, & Kalberg, 2012). Originally published in 1990 for use in 1st through 6th grades, the newest edition extends the assessment to include students in grades 7-9 and established a new form for prekindergarten to kindergarten children (Walker et al., 2014). The SSBD is comprised of two stages. In Stage 1, teachers rank the top five externalizers and the top five internalizers based on a list of example and non-example behaviors. The top three ranked students in each dimension proceed to Stage 2, where they are rated by their teacher on a Critical Events Index (CEI) and Combined Frequency Index (CFI), the latter consisting of the Adaptive and Maladaptive Behavior scales.

The Critical Events Index contains 33 infrequent, but high-risk behaviors that are predictive of current and future behavior problems. Teachers select “Y” or “N” based on the presence or absence of such behaviors as “steals” or “sets fires.” Because of the high risk associated with these items, externalizing students who score 5 or more on the CEI and internalizers who score 4 or more automatically proceed to further observation and assessment, regardless of their scores on the Maladaptive and Adaptive Behavior Scales. Research has demonstrated the sound sensitivity and specificity of the CEI and concurrent validity with the *Achenbach Teacher Report Form* (TRF) and the *Student Risk Screening Scale* (SRSS) teacher and parent ratings (Block-Pedego, Walker, Sevenson, Todis, & Barckley, 1989; Walker et al., 2009)

The Critical Frequency Index (CFI) consists of the Adaptive and Maladaptive behavior scales, comprised of 12 and 11 items respectively. Teachers rate each item on a 5-point Likert

scale ranging from “never” to “frequently”. The CEI, Maladaptive Behavior Scale, and Adaptive Behavior Scale scores are then combined and compared with normed cutoff points to determine whether the student should proceed to further evaluation or intervention (originally called Stage 3 in the first edition, but now included in the *Screening, Identifying, and Monitoring System* [SIMS]; SSBD Administrator’s Guide, pp xiii). The entire process of using the SSBD is designed to take less than one hour to assess an entire class. Over the past 25 years, a host of studies has consistently demonstrated the SSBD’s strong psychometric properties (Kalberg, Lane, & Menzies, 2010; Elliot & Busse, 2004; Epstein & Cullinan, 1998). A particular advantage of the SSBD is that it is one few behavior screeners to assess risk for both internalizing and externalizing disorders, an important feature considering the increase of internalizing disorders during adolescence (Herman, Merrell, & Reinke, 2004). Overall, evidence shows how the strengths of the SSBD make it the preferred screening system for many researchers and practitioners.

Use of the SSBD in Middle Schools

Although the SSBD is considered the gold standard for behavior screeners and has recently been approved for use in middle schools, few studies have been performed to fully validate its use for this population. In 2008, Caldarella, Young, Richardson, Young, and Young conducted a study on the SSBD in two Utah middle and junior high schools. The combined population of the two schools was 2,146 students, 123 of whom were ranked by more than one teacher and advanced to Stage 2 (Caldarella et al., 2008). Their results demonstrated strong internal consistency (.90 for the Maladaptive scale and .84 for the Adaptive scale) and adequate interrater agreement (.44-.75). They also found evidence of concurrent validity, in that Stage 1 successfully differentiated between ranked and non-ranked students on ODRs, attendance, GPA,

the TRF rating scale, and SRSS. Discriminant and convergent validity of internalizers and externalizers was also found with the externalizing and internalizing subscales of the SRSS and the TRF.

A second study by Richardson, Caldarella, Young, Young, and Young (2009) extended this research by replicating their earlier results and further exploring the correlations between the SSBD Stage 2 and teacher, parent, and student ratings. They again found adequate interrater correlations (.44-.61) and strong internal consistency coefficients (.88-.90). They also found modest correlations between the SSBD Stage 2 scales and the ASEBA Parent *Child-Behavior Checklist* (CBCL) and *Youth Self Report* (YSR), along with the SSRS parent form. Together, these two studies provide preliminary evidence for using the SSBD in middle and junior high schools.

Limitations of the SSBD-2 and Universal Screening in Middle Schools

Despite the above evidence, there remain several limitations to the previous research in determining the validity and reliability of the SSBD when used with early adolescents. First, to our knowledge no study has examined the rate of risk detection of the SSBD in identifying EBD in this population. The average middle school class contains 25.5 students (U.S. Department of Education, 2012). Because most middle school teachers have from 4-7 classes a day, they may teach upwards of 100-150 students total. The SSBD allows teachers to nominate only three internalizers and three externalizers, which amounts to roughly 5% of their students proceeding to Stage 2 (as opposed to the roughly 1/4th of students who could be nominated in a self-contained elementary school class; U.S. Department of Education, 2012). Considering that an estimated 20% of students exhibit characteristics of EBD and that the number of disciplinary actions increases in middle school, to achieve sufficient sensitivity levels it may be necessary to

rate more than the top six students (Sugai, Sprague, Horner, & Walker, 2000). As compared to other screeners, multi-gated assessments must be strongly sensitive to prevent at-risk students from being overlooked and missing the opportunity to receive further assessment. In other words, multi-gating approaches should be less concerned about false positives in the first gate (Glover & Albers, 2007; Albers & Kettler, 2014). To investigate the rate of risk detection of the SSBD-2 for this population, more students beyond the top three on each dimension should be rated on the Stage 2 scales in order to evaluate at what point a more appropriate cutoff should be made. Further analyses should compare the rate of risk identification of the SSBD standardization sample with the SSBD in middle school students. Higher rates of risk detection (as well as the rate of risk identification across the rank of different students) could indicate inappropriate cutoff scores for progressing to further evaluation and intervention.

Another limitation is in regards to the logistical concerns of administering the SSBD to a large number of teachers of different subjects. Elementary school students typically have 1 or 2 teachers, with one being designated as a homeroom teacher. Alternatively, middle school students have upwards of 6 to 8 teachers in different academic areas. The SSBD Administrator's Guide is unclear as to how data collection in middle schools might differ from that in elementary schools. Collecting data from all teachers who consider all of their students may be unfeasible to analyze and organize with the limited resources many schools have. In that scenario, multiple teachers might possibly rank and complete the SSBD for the same students, creating a significant amount of paperwork and hassle for school staff. In the case of limited resources, it would be more practical for one teacher to rate each student. One way to implement that procedure is for all the teachers (math, science, etc.) in one period to complete the SSBD for their students only in that class (e.g., homeroom). In this scenario, no student would be ranked by more than one

teacher; however, differences in teacher ratings by subject could result in different rates of risk detection. For example, if social studies teachers are found in general to rate their students lower on the Maladaptive Behavior Scale, their scores may result in fewer students identified at risk. There is no known research on differing ratings of student behavior of teachers across different academic subjects, using the SSBD or other scales. An analysis of possible differences in teacher responses on the SSBD scales and other measures of behavior may provide interesting insight into varying perceptions of student behavior. Because no known studies have examined these factors, data analyses are considered exploratory.

Purpose of Study

The SSBD is a well-established measure for identifying students at risk for EBD. The purpose of this study is to provide further evidence for using the SSBD with middle school populations by investigating its reliability, validity, and rates of risk detection for identifying students at risk for emotional and behavioral disorders. An additional purpose is to investigate the differences in teacher responses and rate of risk identification across academic subjects. This study addressed the following research questions:

1. Are there significant differences between students ranked in the top three in the SSBD Stage 1 nominations (“ranked”) and those in the top five (“non-ranked”)?
2. Is there evidence for concurrent validity of the SSBD Stage 1 nominations with other measures of behavior?
3. Is there evidence of adequate internal consistency in the Stage Two Adaptive and Maladaptive Behavior Scales?
4. Is there evidence of convergent and discriminant validity of the Stage Two Adaptive and Maladaptive Behavior scales with other measures of externalizing/internalizing behavior?

5. How do rates of risk detection in middle school students compare to those in the normative sample?
6. Are there significant differences in the responses and rates of risk identification between teachers of different subjects?

METHOD

Participants and Setting

Sixteen teachers were recruited from two middle schools in Southeastern Louisiana. Four teachers from each science, mathematics, English/language arts, and social studies were included in the sample. Three 6th grade teachers, eight 7th grade teachers, and five 8th grade teachers agreed to participate in the study. Of these teachers, 10 were female and 6 were white. The average teacher age was 33.9 (ranged 21-56) and the mean years in the education field was 9 years (ranged 1-31). The study procedures took place in fall of the 2016 school year.

Table 1.
Teacher Demographic Information

	<i>n</i>	%
Gender		
Female	10	62.50%
Male	6	37.50%
Ethnicity		
Caucasian	12	75%
African American	4	25%
Grade Taught		
6th	3	18.75
7th	8	50%
8th	5	31%
	<i>n</i>	<i>M</i>
Age	16	33.89
Years Taught	16	9.59

Measures

SSBD Stage 1. For the traditional administration of the SSBD Stage 1, teachers are instructed to rank the top five students in their classes on both the internalizing and externalizing behavior dimensions. Teachers are given brief definitions of internalizing and externalizing behaviors as well as examples (“appearing depressed” or “defying the teacher”) and non-

examples (“initiating social interactions with peers” or “following directions”) of behaviors on each dimension. If students appear to exhibit both internalizing and externalizing characteristics, teachers are instructed to rank the student on the dimension that best matches their overall behavior. Students cannot be ranked on both the externalizing and internalizing dimensions. After the top five students on each dimension have been ranked, the top three move on to Stage 2. For the purposes of this study, teachers were instructed to complete the Stage 2 scales for all five students in each dimension. This alteration in the procedure allowed for the exploration of the ratings of students who were ranked beyond the top three on each dimension.

SSBD Stage 2. The SSBD Stage 2 consists of the Critical Events Index (CEI) and the Combined Frequency Index for Adaptive and Maladaptive Behavior (CFI). The CEI is a list of low-frequency but high-risk behaviors on both the internalizing and externalizing dimensions in which the teacher indicates whether the behavior has occurred since the last assessment or the beginning of the school year. Example items include, “vomits after eating”, “damages others’ property”, and “is teased, neglected and/or avoided by peers.” Previous studies of the SSBD in middle schools (Caldarella et al, 2009) divided the CEI into Internalizing and Externalizing subscales, which were used in validity analyses. The present study also uses these subscales. The Adaptive and Maladaptive Behavior subscales on the CFI consist of 12 and 11 items respectively in which teachers rate behaviors on a 5-point Likert scale ranging from “never” to “frequently.” These subscales include such items as “is considerate to the feelings of others” and “tests or challenges teacher-imposed limits.” Students who exceed normed cutoff scores for Stage 2 proceed to the optional Screening, Identification, and Monitoring System (SIMS), which includes academic and social observations and a school records search. Otherwise, identified students progress to intervention or further assessment per the school’s discretion.

Social Skills Improvement System-Rating Scales: Problem Behaviors Scale -

Internalizing/Externalizing Subscales. The SSIS-RS assesses students on three domains: social skills, problem behaviors, and academic competence (Gresham & Elliot, 2008). The Problem Behaviors subscales include 19 items pertaining internalizing and externalizing behaviors. Each item is rated on a 4-point frequency scale from “never” to “almost always” when considering the student’s behavior in the past 2 months.

ODRs and Letter Grades. For each student, teachers provided the current letter grade and number of Office Disciplinary Referrals (ODRs) issued from their classroom.

Procedure

Teachers were recruited from two middle schools in the East Baton Rouge Parish school system. Upon obtaining administrator and teacher consent to participate, each teacher was given a packet including the SSBD, the internalizing/externalizing items of the SSIS-RS Problem Behaviors subscale, and a place to record the student's ODRs and letter grade. They were given a week to complete the packets, which were then retrieved by the experimenter. For their participation, they were granted a "free dress day" and entered into a chance to win \$50.

Analyses

Analyses of SSBD Stage 1. To determine if students who were ranked in the top three internalizers/externalizers (ranked) significantly differed from students who were rated fourth and fifth (non-ranked), independent *t* tests were performed between ranked and non-ranked students from each dimension on ODRs, letter grade, Stage 2 scores, the SSIS-RS internalizing items and externalizing items. If there are no significant behavioral differences between these groups, ranking only the top three students may not be enough to fully capture at-risk students. To further assess the validity of the Stage 1 dimensions, externalizing students were compared

with internalizing students using independent t tests on ODRs, letter grade, Stage 2 scores, and the SSIS-RS internalizing and externalizing items. Further Stage 1 analyses compared teacher responses to behavior measures across academic subjects, using 4x1 between-groups ANOVAs on the previous factors. T -tests using a Bonferonni correction followed up significant ANOVAs. Because there is no prior evidence of between-subject comparisons, these analyses are considered exploratory.

Analyses of SSBD Stage 2. Internal consistency was examined by calculating Cronbach's alpha. Convergent and discriminant validity were calculated using Pearson's r coefficients between the Stage 2 scores and the SSIS-RS internalizing and externalizing items, ODRs, and letter grade (the latter being exploratory). Rate of risk identification was determined by computing the percentage of internalizing and externalizing students who meet cutoff for at-risk status based on the combination of CEI, MBS, and ABS scores. Rates of risk identification were calculated with all teachers combined and across teachers of different subjects.

RESULTS

Descriptive Statistics

Descriptive statistics are provided in Table 2 to summarize data on the SSBD and other measures of behavior ($N = 160$).

Table 2.

Descriptive Statistics for the SSBD Stage 2 Scales and Other Behavior Measures

	<i>M</i>	<i>SD</i>	Range
SSBD Stage 2 CEI	3.89	2.87	0-13
SSBD Stage 2 MBS	25.47	9.98	11-53
SSBD Stage 2 ABS	37.98	9.28	18-57
SSIS-RS Internalizing	12.62	4.36	7-26
SSIS-RS Externalizing	21.99	8.18	12-48
ODRs	.44	.81	0-4
Letter Grade	2.68	1.23	1-5

Results of the SSBD Stage 2 Critical Events Index (CEI) revealed a mean of 3.89 and a standard deviation of 2.87. The number of critical events endorsed ranged from 0 to 13. The SSBD manual suggests that a score of 10 or more is extremely rare. Teachers reported 10 or more critical events for 8 (5%) students in this sample. Results of the Stage 2 Combined Frequency Index (CFI) revealed a mean of 25.47 and standard deviation of 9.98 on the Maladaptive Behavior Scale (MBS), and a mean of 37.98 and standard deviation of 9.28 on the Adaptive Behavior Scale (ABS). Most teachers issued very few ODRs, at less than one per student ($M = .44$; $SD = .81$) and with a range of 0 to 5. When presenting the study materials,

several teachers stated that most behavioral incidents are handled informally, without documented ODRs. Additionally, because the parameters of this study precluded the availability of individual student data, reported ODRs are based on teacher estimates. For these reasons, the number of ODRs may underestimate the frequency of problem behavior in each teacher's classroom. Letter grade was analyzed by coding letter grades by number (F = 1, D = 2, C = 3, B = 4, and A = 5). The mean grade was 2.68 with a standard deviation of 1.23, indicating the average student grade is between a D and a C. Results of the SSIS-RS Internalizing subscale revealed a mean of 12.62 and standard deviation of 4.36, and the SSIS-RS Externalizing subscale produced a mean of 21.99 and standard deviation of 8.18.

Mean Differences Between Stage 1 Ranked vs. Non-ranked Students

To assess for group mean differences between ranked and non-ranked students, independent t-tests were performed to assess differences in SSBD Stage 2 scores, ODRs, grade, and the SSIS-RS internalizing and externalizing. For the purpose of these analyses, "ranked" students were those ranked in the top 3 internalizers and externalizers according to provided descriptions and "unranked" students were those ranked as 4th and 5th internalizers and externalizers. The groups are thus named because in the standard administration of the SSBD, only the top three students from each dimension would be nominated to complete the Stage 2 scales. Results yielded significant differences between groups on ODRs and the Critical Events Index. In some cases, the assumption of homogeneity of variances was violated, according to Levene's test. In these cases, adjusted t-statistics were reported.

Table 3.

Group Mean Comparisons Between Ranked and Non-Ranked Students

Scale	Ranked N=96	Non-Ranked N=64	Mean Difference	Standard Error of Difference	<i>t</i>
ODRS	.56	.26	.296	.117	2.54*†
Grade	2.77	2.55	.278	.215	1.09
SSBD Stage 2					
CEI	4.26	3.34	.925	.43	2.15*
MBS	26.5	23.95	2.56	1.52	1.68
ABS	37.73	36.64	.56	1.45	.387†
SSIS-RS					
Internalizing	12.88	12.23	.652	.703	.93
Externalizing	22.8	20.92	1.89	1.23	1.52

* $p < .05$

Note. †Independent samples t-test with Bonferroni correction, in which Levene's test for equality of variances showed a violation of the assumption of homogeneity of variances. Adjusted *t* statistic reported.

Teachers reported more ODRs on average for ranked students ($M = .56$, $SD = .93$), than non-ranked students ($M = .26$, $SD = .54$), $t(158) = 2.54$, $p < .05$, $d = .39$. Teachers also endorsed more critical events for ranked students ($M = 4.26$, $SD = 3.19$) than non-ranked students ($M = 3.34$, $SD = 2.24$), $t(158) = 2.15$, $p < .05$, $d = .32$. Significant group differences were not found for grade, $t(158) = 1.09$, $p = .28$, $d = .18$, ABS scores $t(157) = .37$, $p = .7$, $d = .06$, or MBS scores, $t(158) = 1.59$, $p = .11$, $d = .26$. Furthermore, scores on the SSIS-RS Internalizing subscale, $t(158) = .97$, $p = .36$, $d = .16$, or the SSIS-RS Externalizing subscale, $t(157) = .162$, $p = .13$, $d = .25$.

These analyses provide evidence that ranked students do not significantly differ from non-ranked

students on most measures, indicating that students not meeting the nomination criteria for the standard administration of the SSBD (i.e., the top three externalizers and internalizers) nevertheless exhibit significant behavior problems.

Table 4.

Group Comparisons Between Externalizing and Internalizing Students

Scale	Externalizing N=80	Internalizing N=80	Mean Difference	Standard Error of Difference	<i>t</i>
ODRS	.74	.14	.6	.12	5.06*†
Grade	2.45	2.91	-.46	.215	-2.43*
SSBD Stage 2					
CEI	4.63	3.15	1.46	.44	3.35*
MBS	30.36	20.49	9.96	1.52	7.23*
ABS	33.53	40.46	-6.92	1.37	-5.04*
SSIS-RS					
Internalizing	11.19	14.08	-2.89	.66	-4.39*
Externalizing	26.54	17.39	9.15	1.08	8.48*

* $p < .01$

Note. †Independent samples t-test with Bonferroni correction, in which Levene's test for equality of variances showed a violation of the assumption of homogeneity of variances. Adjusted *t* statistic reported.

Mean Differences Between Stage 1 Internalizing and Externalizing Students

To further assess for concurrent validity of the Stage 1 rankings, independent t-tests were conducted between internalizing and externalizing students on the Stage 2 scales and other measures of behavior. As above, the assumption of homogeneity of variance was violated in several cases according to Levene's statistic. In those cases, an adjusted *t*-statistic was reported.

Results are presented in Table 4. Statistically significant differences between identified internalizing and externalizing students were found on all variables. Teachers issued more ODRs for externalizing students ($M = .74, SD = .96$) than internalizing students ($M = .14, SD = .44$), $t(158) = 5.06, p < .001, d = .8$, as expected given the nature of most ODRs (i.e., rule infractions). Furthermore, teachers endorsed more critical events on the CEI scale for externalizing students ($M = 4.6, SD = 3.09$) than internalizing students ($M = 3.15, SD = 2.44$), $t(158) = 3.35, p = .001, d = .52$. Significant positive differences were also found for the MBS $t(156) = 7.23, p < .001, d = 1.15$. The SSIS-RS Internalizing and Externalizing subscales provided additional evidence for concurrent validity for the Stage 1 nominations. On the SSIS-RS Internalizing subscales, internalizing students scored higher ($M = 14.08, SD = 4.26$) than externalizing students ($M = 11.19, SD = 4.0$), $t(156) = -4.39, p < .001, d = .7$. Conversely, nominated externalizing students scored significantly higher ($M = 26.5, SD = 6.71$) on the SSIS-RS Externalizing subscale than internalizing students ($M = 17.39, SD = 6.86$), $t(158) = 8.49, p < .001, d = 1.34$. Overall, these results provide evidence as to concurrent validity of the Stage One rankings, by establishing significant comparisons of identified internalizing and externalizing students with other measures of corresponding behaviors.

Reliability of the SSBD-2 Adaptive Behavior and Maladaptive Behavior Scales

The consistency of item responses on the MBS and ABS was assessed by calculating Cronbach's coefficient alpha. Reynolds and Livingston (2015) suggest that, given the low-stakes decisions of screeners, reliability coefficients be at least .70 (2015). In the present study, teacher ratings on both scales yielded reliability estimates exceeding that standard. Teacher responses on both the MBS and ABS generated Cronbach's alpha of .89.

Convergent and Discriminant Validity of the Stage 2 Scales

Convergent and discriminant validity was assessed by correlating the SSBD Stage 2 scales in relation to other established measures of behavior. For the purpose of these analyses, the Critical Events Index was separated into two subscales representing internalizing and externalizing items. A prior study (Caldarella et al., 2008) developed these subscales by selecting items that significantly discriminated between internalizing and externalizing students on the Stage 1 rankings. Theoretically, scores on the CEI Internalizing and Externalizing subscales should correlate significantly with other measures of internalizing and externalizing behavior. Validity estimates were determined by calculating Pearson's r between the SSBD Stage 2 scales and the SSIS-RS Internalizing and Externalizing subscales, ODRs, and letter grade. The strength of correlations were interpreted based on Cohen's (1997) recommendations. Correlations more than .5 are considered large, .30-.49 are moderate, and below .3 are small.

Table 5.

Correlations Between the Systematic Screening for Behavior Disorders and the SSIS-RS, ODRs, and Letter Grade

	SSIS Internalizing	SSIS Externalizing	ODRs	Letter Grade
CEI Internalizing	.615**	-.186*	-.142	.027
CEI Externalizing	-.208**	.803**	.418**	-.098
MBS	-.083	.895**	.345**	-.065
ABS	-.064	-.673**	-.297**	.219*

* $p < .05$. ** $p < .01$

Table 5 represents findings of the current study and provides evidence for adequate convergent and discriminant validity of the Stage 2 measures. Convergent validity is evidenced by strong positive correlations between the SSIS-RS Internalizing and CEI Internalizing

subscales ($r = .615$) and between the SSIS-RS Externalizing and CEI Externalizing subscales ($r = .803$). Conversely, discriminant validity is demonstrated through negative correlations between the SSIS Internalizing and CEI Externalizing subscales ($r = -.208$) and the SSIS Externalizing and CEI Internalizing subscales ($r = -.186$), although these correlations are weaker.

Other correlations between the Stage 2 scales included correlations between the SSIS-RS Externalizing scales and other behavior measures. The SSIS-RS Externalizing subscale yielded strong correlations with not only the CEI Externalizing subscale ($r = .803$), but also the MBS ($r = .895$) and the ABS ($r = -.673$). Number of ODRs was found to be moderately correlated with externalizing critical events ($r = .418$) and responses on the MBS ($r = .345$), as would be expected based on the externalizing nature of many rule infractions and maladaptive behaviors. Letter grade had a small correlation with ABS ($r = .219$), indicating that grades marginally increase with adaptive behaviors.

Risk Identification

Decision rules regarding risk identification are based on the results of the Stage 2 scores, including the number of endorsed critical events on the CEI and scores on the Maladaptive and Adaptive Behavior scales. On the SSBD 1st Edition, students who qualified as at risk based on Stage 2 criteria would automatically proceed to Stage 3 measures (including observations and school archival records search); however, the SSBD 2nd Edition removed Stage 3, which is now encompassed under the Screening, Identification, and Monitoring System (SIMS). Users of the SSBD have the option of completing Stage 3 measures or starting immediate intervention or referral. Results of differences in risk detection between the SSBD Administrator's Guide and in the current study are presented below.

Table 6.

Rate of Risk Identification of the SSBD Manual and the Present Study

Percentage Identified At-Risk	SSBD Guide	Present Study
Internalizers	29.7%	46.8%
Externalizers	33.8%	46.3%

Risk criteria differ between students identified as internalizing or externalizing on Stage 1 nominations and are listed in the SSBD Administrator's Guide (pp 79-80) and below.

Table 7.

Decision Rules for Externalizing Risk.

Critical Events	Adaptive High=Lower Risk	Maladaptive High =Greater Risk	At Risk?
0			No
1-4	31 or more		No
	30 or less	34 or less	No
		35 or more	Yes
5 or more			Yes

Borrowed from *SSBD Administrator's Guide* (pp 79)

According to the criteria listed above, 46.3% ($N=37$) of externalizing students were identified as at risk for developing emotional and behavioral disorders. This compares to the 33.8% of ranked externalizers in the SSBD standardization sample who met criteria (Administrator's Manual, pp 93). Risk status was further analyzed by Stage 1 rank (i.e., 1-5). Results revealed that 13.75% of externalizing students ($N=11$) ranked 4th or 5th qualified as at risk. Based on standard administration of the Stage 1 procedures, Stage 2 measures would only be completed on the top three externalizers. When using that procedure, these 11 students would not have been identified, despite meeting criteria.

Table 8.

Decision Rules for Internalizing Risk.

Critical Events	Adaptive High=Lower Risk	Maladaptive High =Greater Risk	At Risk?
0			No
1-3	42 or more		No
	41 or less	18 or less	No
		19 or more	Yes
4 or more			Yes

Borrowed from the *SSBD Administrator's Guide* (pp 80)

Similar analyses were performed for internalizing students on Stage 1. On the above criteria, 46.3% (N=37) of internalizing students were identified as at risk. This compares to the 29.7% of internalizers who qualified based on standardization norms (Administrator's Manual, page 94). When analyzed by rank, 17.5% (N=14) were not ranked in the top three, and would thus not have been identified using standard SSBD administration procedures.

Group Mean Differences Between Teachers Responses Across Subjects

To assess differences in teacher responses across subjects, a series of 4x1 one-way ANOVAs were performed. Results of these analyses are provided in Table 9. Significant differences were followed up by t-tests with Bonferroni corrections.

Table 9.

Group Mean Differences Across Academic Subjects.

	ELA		Math		Science		Social Studies		F-value
	M	SD	M	SD	M	SD	M	SD	
SSBD Stage 2									
CEI Total	3.2	2.59	5.25	3.38	4.35	2.82	2.78	1.94	6.65**†
CEI Internalizing	.83	1.01	1.23	.95	.78	.92	.86	.99	1.78
CEI Externalizing	1.3	1.48	2.85	2.38	2.05	1.98	1.45	1.57	5.59**†
MBS	25.74	11.92	28.56	9.62	27.69	8.31	20.05	7.51	6.44**
ABS	36.95	9.53	35.23	8.37	36.68	9.68	39.1	9.44	1.18†

**p < .01

Note. †Independent samples t-test with Bonferroni correction, in which Levene's test for equality of variances showed a violation of the assumption of homogeneity of variances. Adjusted F statistic reported.

The analyses revealed significant differences between teacher ratings on the CEI, $F(159) = 6.65, p < .01$. Follow-up analyses revealed math teachers tended to endorse more critical events ($M=5.25, SD = 3.38$) than do ELA teachers ($M=3.2, SD=2.59$) or social studies teachers ($M=2.82, SD=1.94$). Similar differences were also found particularly on the externalizing items of the CEI, $F(159) = 5.59, p < .01$. Scores were found to be higher for math teachers ($M=2.85, SD=2.38$) than ELA teachers ($M=1.3, SD=1.48$) or social studies teachers ($M=1.45, SD=1.57$). Finally, significant differences were found between teacher responses on the MBS, $F(157)=6.44, p < .01$. In particular, social studies ($M=20.51, SD=7.51$) teachers rated their students' behavior less severe than math teachers ($M=28.56, SD=9.62$) or science teachers ($M=27.69, SD=8.31$). These results align with differing rates of risk identification across teachers, where social studies identified fewer students.

Table 10.

Rate of Risk Identification Across Subjects.

Percentage Identified At-Risk	ELA	Math	Science	Social Studies
Internalizers	40%	60%	65%	20%
Externalizers	50%	55%	50%	30%

DISCUSSION

Emotional and behavioral difficulties often have a detrimental impact on student academic performance and social development, frequently resulting in negative short- and long-term outcomes. Universal screening is essential for identifying students at risk for EBD and targeting them for intervention to prevent or mitigate the adverse consequences of these disorders. Middle school students are especially at risk for developing emotional and behavioral problems, but few screeners are available for assessment of this population. The present sought to replicate previous research on the use of the *Systematic Screening for Behavior Disorders* in middle school settings, as well as explore differences in teacher responses across academic subjects. The specific research questions addressed (a) differences between ranked and non-ranked students on the SSBD Stage 1 nominations (b) evidence of concurrent validity of the Stage 1 nominations with other measures of behavior (c) evidence of adequate reliability of the SSBD Stage 2 scales (d) evidence for convergent and discriminant validity of the SSBD Stage 2 scales and (e) differences in teacher responses and risk identification across academic subjects.

The first research question addressed differences on measures of behavior between students ranked in the top three internalizers/externalizers and those ranked in the top five on the Stage 1 nominations. The results of these analyses indicate that ranked and non-ranked students display similar scores on most measures of behavior. Two exceptions to this rule are ODRs and CEI scores. Both the number of ODRs and critical events were significantly higher in ranked students than non-ranked students. However, average ODRs were low for both ranked ($M=.56$) and non-ranked ($M=.26$) groups, significantly lower than the number of ODRs in other studies using the SSBD (Caldarella et al., 2009). A frequency distribution revealed that most students received no ODRs at all and the majority of ODRs were issued to a small number of select

students, suggesting that these few students may have particularly severe behavior in comparison to their peers. Anecdotally, several teachers reported that most behavioral incidents are handled within the classroom and do not result in official documentation of infractions. Furthermore, number of ODRs were gathered by teacher report rather than from school records. The low total ODRs, possible fallibility of teacher report, and the distribution of ODRs across students may have resulted in the observed data.

On the CEI, while teachers similarly endorsed significantly more critical events for ranked students ($M=4.26$) than non-ranked students ($M=3.34$), the means of both groups signify moderate to high levels of risk. According to the SSBD risk identification decision rules, CEI scores between 1 and 4 (for externalizing students) and between 1 and 3 (for internalizing students) indicate a student is at moderate risk of developing emotional and behavioral difficulties, while even higher CEI scores suggest high risk. Although ranked and non-ranked students significantly differed on number of critical events endorsed, their scores do not represent significantly different levels of risk based on cutoff criteria. These results indicate that some at-risk but non-ranked students would not have been included in the Stage 2 procedures based on standard administration of the SSBD and thus would not have been identified by this measure.

The second research question explored levels of concurrent validity of the SSBD Stage 1 nominations with other indicators of internalizing and externalizing behaviors. The results provide evidence for the concurrent validity of the Stage 1 internalizing and externalizing nominations across all other measures of behavior. As expected, students nominated as having externalizing behavior scored higher on measures of externalizing behavior, such as the SSIS-RS Externalizing Subscale and the Maladaptive Behavior Scale (which is largely comprised of

externalizing-type items, such as “creates a disturbance during class activities.”), and ODRs. In contrast, students nominated as internalizers scored higher on the SSIS-RS Internalizing subscale and Adaptive Behavior Scale, and also demonstrated increased academic performance compared to their counterparts. Unlike externalizing behaviors, internalizing behaviors, such as shyness or low activity level, generally do not preclude adaptive behaviors such as following classroom rules or completing work, which possibly produced the observed higher ABS scores and letter grade when compared to students with externalizing behaviors. In addition, results demonstrated that nominated internalizing students received significantly lower CEI scores than nominated externalizing students. Regarding these divergent scores, internalizing-type items on the CEI (like internalizing characteristics more generally) may not be readily visible to teachers. For example, a teacher would not necessarily be aware of a student’s nightmares (CEI Item 12) or anhedonia (CEI Item 33) without being informed by that student. Reports of these behaviors may be even less frequent given the limited time a middle school teacher spends with each of his/her students per day, roughly 50 minutes to an hour per day. Overall, observed differences in scores between nominated internalizing and externalizing students provide evidence for concurrent validity of the Stage 1 nominations.

The third research question examined the reliability of the SSBD as assessed by internal consistency estimates. Both the ABS and MBS achieved internal consistency estimates of $\alpha = .89$, which are consistent with reliability coefficients found in other studies of the SSBD with middle school students (Cronbach’s $\alpha = .83-.90$; Caldarella et al., 2009) as well as in the SSBD standardization sample (Cronbach’s $\alpha = .92 - .94$; SSBD Manual, pp 32). Because screeners generally result in low-stakes decisions (Reynolds & Livingston, 2014), the suggested

threshold for reliability coefficients is .70. The internal consistency estimates for both the ABS and MBS easily meet this threshold.

The fourth research question addressed measures of convergent and discriminant validity of the SSBD Stage 2 scores with other measures of behavior. Strong evidence of convergent validity was demonstrated through positive correlations between the SSIS-RS Internalizing items and the internalizing items on the CEI ($r = .615$). Conversely, scores on the externalizing items on the CEI were strongly correlated with the SSIS-RS Externalizing items ($r = .803$) and moderately correlated with ODRs ($r = .418$). The latter correlation is to be expected given the externalizing nature of ODRs, which typically cover rule infractions such as disruptive behavior, and the likelihood that many events on the CEI would result in ODRs, particularly high risk events such as setting fires or assaulting other students. Similarly, the MBS also demonstrated moderate to strong correlations with the SSIS-RS Externalizing items ($r = .895$) and ODRs ($r = .345$). Overall, these results provide evidence that the items on the Stage 2 scales represent the same constructs as similar measures of behavior. Evidence for discriminant validity was supported by significant negative correlations between the CEI Internalizing and the SSIS-RS Externalizing subscales. While the relationship between the CEI Internalizing subscale and ODRs trended in the negative direction, this correlation was not significant. This finding differs from previous studies; however, the difference in the current study may be due to the particularly low rate of ODRs for internalizing students.

The fifth research question explored rates of risk identification of the SSBD for internalizing and externalizing students. Just under 50% of students were determined to be at risk based on decision rules enumerated in the SSBD Administrator's Guide. Based on these criteria, the rate of risk identification is higher for middle school students than elementary school

students. Several factors may explain the difference in rates of risk identification. First, the CEI has not been examined at the item-level in middle school populations. It may be that some critical events constitute greater risk for elementary school students than middle school students, and thus be inappropriate indicators of risk for adolescents. For example, “Uses obscene language or swears” (CEI 21) is typical for young teenagers (as evidenced by five minutes in a middle school hallway) and may well be less severe of a behavior in this population than with younger children. The actual risk related to critical events is particularly important in the SSBD given the weight of the CEI for risk identification. Externalizing students require 5 or more critical events, while internalizing students only require 4 to automatically meet cutoff scores. In the current study, 28 internalizers and 34 externalizers met criteria based on number of critical events alone, representing a total of 83.8% percent of all students identified as at risk. Further research is needed to examine the association of different critical events with future risk, and determine whether some items should be dropped from the scale for this population or if cutoff scores should be adjusted to improve specificity rates (if such research concludes an adjustment is needed).

In reference to sensitivity, several students ($N=25$) in ranks four and five (as nominated in Stage 1) met or exceeded cutoff scores, but would not have been identified based on the standard administration of the SSBD, under which only the top three ranked students are rated on Stage 2 measures. Given the extensive negative outcomes associated with EBD, measure sensitivity is vital to the successful detection of students at risk. This is particularly important for gated screeners, because students who fail to pass the first gate automatically do not qualify for risk status (Glover & Albers, 2007). In the current study, practicality limited the number of rated students to the top five ranked for each dimension. Therefore, it is impossible to know how many

students past the top five would qualify for further intervention or evaluation. Future research is needed to examine true sensitivity and specificity rates for the SSBD in middle schools.

A sixth research question addressed possible differences in teacher perceptions of student behavior across subjects as measured by responses on the and risk identification rates. Regarding ratings on the SSBD Stage 2 scales and other measures of behavior, small differences emerged between subjects but they did not appear to represent overarching patterns in teachers' responses or perceptions of students. Math teachers endorsed more critical events than did other teachers, but did not differ from other teachers in their ratings on the MBS. Conversely, social studies teachers' scores on the MBS were lower than other teachers, but there were no differences across teachers on the ABS. In general, teacher perceptions of students did not seem to significantly differ based on the SSBD. The results could indicate that student behavior tends to be stable across classes, or that teachers have consistent interpretations of items on the SSBD. Future research should use other methods to study differences in perceptions of behavior across subjects, such as interrater correlations and behavioral observations.

In terms of risk identification, social studies teachers identified fewer students as at-risk than teachers of other subjects. This corresponds to their slightly lower scores on some other measures of behavior. However, there were only four teachers in each subject group, resulting in 40 students per subject (20 internalizing students and 20 externalizing students). With such a small sample size, it is possible one teacher endorsed fewer critical events or provided lower responses on the MBS and ABS, bringing the entire rate of risk identification down for that subject. Future studies with more teachers could provide more information about differences in risk detection across different academic subjects.

Implications for Practice

The current study offers important evidence for the reliability and validity of the SSBD for identifying students at risk for emotional and behavioral disorders in middle school. Prior research has clearly established the negative outcomes associated with EBD and its emergence during the early adolescent years, but few screening devices exist for use with this population. The SSBD offers schools an efficient and effective measure for identifying and targeting students susceptible to developing emotional and behavioral difficulties.

Unlike many other screeners, the multi-gated structure of the SSBD has the advantage of allowing teachers to consider all of their students for nomination, while only requiring them to complete the rating scales for the few deemed particularly at risk. Some screeners are intended to be completed by teachers for all their students, which is impractical for middle school teachers teaching upwards of 100 students per day. One implication of the current study is that teachers may need to nominate and rate more students to fully capture all students at risk. Students ranked in the top three did not differ significantly than non-ranked students on most markers of risk. Further, total scores on the CEI, MBS, and ABS indicate that ranked and non-ranked students yielded similar probability of meeting risk identification cutoff scores. When using the SSBD, schools should consider how many students should be nominated and rated by each teacher to achieve ideal levels of risk detection in their schools.

A further implication of this study is the success with which the SSBD identified internalizing students as at risk. Anecdotally, several teachers reported that, while identifying externalizing students was easy (and occasionally difficult to narrow down), recognizing internalizing students presented a greater challenge. However, the current study indicates that internalizing and externalizing students were similarly likely to be identified as at risk by the

SSBD Stage 2 totals. These results highlight the importance of screening specifically for internalizing behaviors. Because internalizing behaviors such as depression, anxiety, and somatic complaints rarely disrupt the classroom, they are often overlooked by teachers (Lane et al., 2007). Without explicitly requiring teachers to consider signs of internalizing disorders, emotionally vulnerable students may not receive the services they need. The SSBD provides such students with the opportunity to be evaluated and identified as at risk for developing more serious symptoms and behaviors.

Limitations

While this study provides evidence as to the usefulness of the SSBD for use with middle school populations and across different teachers, there remain several limitations. A significant limitation of this study is the lack of individual student derived from school records. Practicality prevented the researcher from acquiring parent and child assent for the number of participants needed for this study. As an alternative, parent consent was bypassed by not collecting identifiable student data, instead relying on teacher report of ODRs and letter grade. This method limits the amount of information obtained and available for analysis. Factors such as race, gender, socioeconomic status, and overall GPA could potentially interact with Stage 1 nominations and Stage 2 ratings in unforeseen ways. In previous studies using the SSBD in middle schools, most students nominated were white (84 – 95% of students) and male (73 – 80% of students; Caldarella et al., 2008; Richardson et al, 2009). Extensive research has demonstrated that African American students are overidentified for special education services, including the classification of Emotional Disturbance (Lane et al., 2010). Lack of race data in this study precludes analysis into potential differences SSBD Stage 1 rankings, Stage 2 measure scores, and risk identification status between white and African American students, as well as the impact of

other student factors. The only individual student data collected were teacher report of the student's letter grade and number of ODRs issued from their class. The number of total ODRs across classes and numerical GPA may better predict students' nominations and scores on the SSBD and risk-identification status.

Despite the lack of individual student data gathered in this study, however, bypassing parent consent circumvents some of the limitations inherent in research on behavior screeners. When parent consent is required, large portions of the student population are excluded from data collection and analysis, likely including some students potentially at risk for EBD. In previous studies of the SSBD, consent forms were distributed to parents for students nominated by their teachers through Stage 1 rankings and only about 50% of consent forms were returned (Caldarella et al., 2008; Richardson et al., 2009). Excluding half of the ranked students from data collection could affect the data regarding validity and reliability estimates of the SSBD and related behavioral measures, as well as influence analyses on sensitivity and specificity in risk detection. While the current study was unable to gather student data or additional data concerning parent and student factors (such as the Achenbach scales), it did include scores from all students nominated on Stage 1 of the SSBD.

While results of the current study provide evidence for internal consistency, other reliability analyses were not performed. Without identifying students, conducting these analyses would be difficult or impossible. Test-retest reliability could be determined by administering the SSBD measures several weeks apart. However, retesting would require teachers to remember which students they nominated and their rankings. For confidentiality purposes, teachers were strongly discouraged to write down student names, even on their own paper. Teacher recall may be unreliable without recording students' identities. Additionally, interrater reliability depends on

correlating teacher ratings for the same students. Matching students between teachers again requires students' identities. Given the lack of identifying data, it is possible some students in the current study may have been nominated by more than one teacher, which would decrease variability in the data. However, given the number of students each teacher teaches, the distribution of grade levels and subjects across schools, and the size of the schools themselves, the chance two teachers selected the same student is not likely.

In terms of examining teacher perceptions across subjects, the sample size of teachers was small in this study. Only four teachers of each subject participated, resulting in 20 internalizing students and 20 externalizing students rated for each class. A larger sample of teachers may demonstrate significant differences in responses between academic classes.

Future Directions

Future research should analyze longitudinal data regarding outcomes for students identified by the SSBD. Students identified as at-risk would be predicted to have poorer outcomes based on other markers of behavior. For example, higher rates of criminal activity, drug use, truancy, or dropout would be expected of at risk students (particularly those nominated for externalizing characteristics) when compared to non-identified students. Further, internalizing risk status may predict future diagnoses of related behaviors such as depression or anxiety. More proximal outcomes measures may include ODRs and GPA at the end of the semester/school year, referral for special education evaluation or services, or suspensions or expulsion. Because the risk identification decision rules are currently normed on elementary student data, predictive validity analyses are particularly important when evaluating the SSBD for use in middle schools. As previously stated, certain items on the CEI or scores on the Stage 2

scales may represent greater risk for elementary students than middle school students, thus increasing the false positive rate for identifying true risk status.

While this study examined the reliability and validity of the SSBD, it did not assess usability and acceptability by teachers. While the SSBD does have the advantage of being less time-consuming than other measures (i.e., completing Stage 2 scales only for select students), teachers could possibly prefer other screening instruments, or dislike screening students at all. A recent study found that only 17% of teachers used behavioral screeners in their school (Byrd & Erickson, 2017). The use of any EBD screener may feel onerous to teachers. Given the documented importance of early identification, comparisons of different behavioral screeners in acceptability and perceived usability may prove informative when comparing measures.

Finally, future research can examine the SSBD Stage 2 scales at the item level. While the division of the CEI into the CEI Internalizing and CEI Externalizing scales has been used in previous studies of the SSBD in middle schools, exploration into the psychometric properties of these subscales has yet to be explored. Furthermore, certain items on the ABS and MBS may not be developmentally appropriate for middle school students. For example, Item 2 of the MBS concerns behavior at recess, which is generally irrelevant in middle school. An in-depth item analysis could provide valuable information as to the appropriateness and validity of items on the Stage 2 measures for use with this population and their relationship with risk identification criteria.

Conclusion

The current study provides evidence for the reliability and validity of the SSBD for identifying students at risk for emotional and behavioral disorders, as well as explores possible differences in teacher responses across academic subjects. Emotional and behavioral screening at

the middle school level is widely under-researched and underutilized (Lane et al., 2007), despite the increased risk of developing emotional difficulties at this age level (Eccles et al., 1991).

These results suggest the SSBD can successfully identify emotionally vulnerable students and target them for intervention or further evaluation. Potentially, the introduction of the SSBD or other screening measures in middle schools would serve to prevent or mitigate the negative outcomes associated with emotional and behavioral problems.

REFERENCES

- Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin*, 85, 1275–1301.
- Albers, C. A., Glover, T. A., & Kratochwill, T. R. (2007). Introduction to the special series: How can universal screening enhance educational and mental health outcomes? *Journal of School Psychology*, 45, 113-116.
- Albers, C. A., & Kettler, R. J. (2014). Best practices in universal screening. In P. L. Harrison, & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 121-131). Bethesda, MD: National Association of School Psychologists.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council for Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: AERA
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Byrd, Shelby. M., & Erickson, M. R. (2017). *Adherence to best practices in grade-level team meetings*. Poster presented at the annual conference of National Association of School Psychologists, San Antonio, TX
- Caldarella, P., Young, E. L, Richardson, M. J., Young, B. J., & Young, K. R. (2008). Validation of the Systematic Screening for Behavior Disorders in middle and junior high school. *Journal of Emotional and Behavioral Disorders*, 16(2), 105-117.
- Carter, A. S., Briggs-Gowan, M., & Davis, N. O. (2004). Assessment of young children's social emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry*, 45, 109–134.
- Cook, B. G., Landrum, T. J., Tankersley, M., & Kauffman, J. M. (2003). Bringing research to bear on practice: effecting evidence-based instruction for students with emotional or behavioral disorders. *Education and Treatment of Children*, 26, 345-361.
- Cross, L. H., & Billingsley, B. (1994). Testing a model of special educators' intent to stay in teaching. *Exceptional Children*, 60(5), 411-421.
- Cullinan, D. (2002). *Students with emotional and behavior disorders: An introduction for teachers and other helping professionals* (pp. 23). Upper Saddle River, NJ: Pearson Education, Inc.

- Davis, M., & Vander Stoep, A. (1997). The transition to adulthood for youth who have serious emotional disturbance: Developmental transition and young adult outcomes. *The Journal of Mental Health Administration*, 24, 400–427.
- Dowdy, E., Doane, K., Eklund, K., & Dever, B. V. (2011). A comparison of teacher nomination and screening to identify behavioral and emotional risk within a sample of underrepresented students. *Journal of Emotional and Behavioral Disorders*, 21, 127-137.
- Drummond, K. (1994). *The Student Risk Screening Scale (SRSS)*. Grants Pass, OR: Josephine County Mental Health Program.
- Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to early adolescents? The impact of educational contexts on early adolescents. *American Journal of Education*, 99(4), 521-542.
- Elliot, S. N. & Gresham, F. M. (2007). *Social Skills Improvement System – Performance Screening Guide*. San Antonio, TX: Pearson Assessments.
- Epstein, M.H., Nordness, P. D., Cullinan, D., & Hertzog, M. (2002). Scale for Assessing Emotional Disturbance: Long-term test-retest reliability and convergent validity with kindergarten and first-grade students. *Remedial and Special Education*, 23, 141–148.
- Glover, T. A. & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology*, 4, 117-135.
- Gresham, F. M., & Kern, L. (2004). Internalizing behavior problems in children and adolescents. In R. B. Rutherford, Jr., M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 262 – 281). New York: Guilford.
- Gutman, L. M., & Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African American students during the middle school transition. *Journal of Youth and Adolescence*, 29(2), 223-248.
- Henderson, K., Klein, S., Gonzales, B., & Bradley, R. (2005). Teachers of children with emotional disturbance: A national look at preparation, teaching conditions, and practices. *Behavioral Disorders*, 31(1), 6-17.
- Holas, I., & Huston, A. C. (2012). Are middle schools harmful? The role of transition timing, classroom quality, and school characteristics. *Journal of Youth and Adolescence*, 41, 333-345.
- Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 *et seq.* (2004) (reauthorization of Individuals with Disabilities Act of 1990)
- Kalberg, J. R., Lane, K. L., & Menzies, H. R. (2010). Using Systematic Screening Procedures to Identify Students Who are Nonresponsive to Primary Prevention Efforts: Integrating

- Academic and Behavioral Measures. *Education and Treatment of Children*, 33(4), 561-584.
- Kamphaus, R. W., DiStefano, C., Dowdy, E., Eklund, K., & Dunn, A. R. (2010). Determining the presence of a problem: Comparing two approaches for detecting youth behavioral risk. *School Psychology Review*, 39, 395-407.
- Kauffman, J. M. (1999). How we prevent the prevention of emotional and behavioral disorders. *Exceptional Children*, 65(4), 448-468.
- Kauffman, J. M. (2001). *Characteristics of emotional and behavioral disorders of children and youth* (7th ed.). Upper Saddle River, NJ: Prentice-Hall Inc.
- Kingslake, B. (1983). The predictive (in)accuracy of on-entry to school screening procedures when used to anticipate learning difficulties. *British Journal of Special Education*, 1, 23-26.
- Kratochwill, T. R., Albers, C. A., & Shernoff, E. (2004). School-based interventions. *Child and Adolescent Psychiatric Clinics of North America*, 13, 885-903.
- 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(4), 211-224.
- Lane, K. L., Kalberg, J. R., Lambert, E. W., Crnabori, M., & Bruhn, A. L. (2010). A comparison of universal screening tools for emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 18(2), 100-112.
- 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, Oakes, W. P., & Kalberg, J. R. (2012). *Systematic screenings of behavior to support instruction: From preschool to high school*. New York, NY: Guilford Press
- Lane, K. L., Oakes, W. P., Menzies, H. M., & Germer, K. A. (2014) Screening and identification approaches for detecting students at risk. In H. M. Walker, & F. M. Gresham (Eds.), *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools* (pp. 129-151). New York, NY: The Guilford Press.
- Lane, K. L., Wehby, J., Robertson, E. J., & Rogers, L. (2007). How do different types of high school students respond to positive behavior support programs? Characteristics and responsiveness of teacher-identified students. *Journal of Emotional and Behavioral Disorders*, 15, 3-20

- Loeber, R. (1991) Antisocial behavior: More enduring than changeable? *Journal of the American Academy of Child and Adolescent Psychiatry*, 30(3), 393-397.
- 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.
- Merikangas, K. R., Jian-ping, Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L.,...Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Study-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(10), 980-989.
- Messick, S. (1987). Validity. *ETS Research Report Series*, 2, PUBLISHED ONLINE.
- Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Student teacher relations and attitudes towards mathematics before and after the transition to junior high school. *Child Development*, 60, 981-992
- Midgley, C., Middleton, M. J., Gheen, M. H., & Kumar, R. (2002). Stage-environment fit revisited: A goal theory approach to examining school transitions. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 109–142). Mahwah, NJ: Lawrence Erlbaum.
- Moffitt, T. E. (1993). Adolescent-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4), 674-701.
- Nelson, J. R., Benner, G. J., Lane, K. L., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children*, 71, 59–73.
- Nelson, J. R., Benner, G. J., Reid, R. C., Epstein, M. H., & Currin, D. (2002). The convergent validity of office disciplinary referrals with the CBCL-TRF. *Journal of Emotional and Behavioral Disorders*, 10(3), 181-188.
- Phillips, J. L. (1982). *Statistical thinking* (2nd ed.). San Francisco: Freeman.
- Richardson, M. J., Caldarella, P., Young, B. J., Young, E. L., & Young, K. R. (2009). Further validation of the Systematic Screening for Behavior Disorders in middle and junior high school. *Psychology in the Schools*, 46(7), 605-615.
- Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology*, 45, 193-223.
- Soodak, L. C., Podell, D. M. & Lehman, L. R. (1998). Teacher, student, and school attributes as predictors of teachers' responses to inclusion. *The Journal of Special Education*, 31(4), 480-497.

- 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.
- Sugei, G. & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*(2), 245-259.
- Taylor, R. L. (2000). *Assessment of exceptional students: educational and psychological procedures (5th Ed.)*. Needham Heights, MA: Allyn & Bacon. LOOK AT THIS ONE
- Trout, A. K., Nordness, P. D., Pierce, C. D., & Epstein, M. H. (2003). Research on the academic status of children with emotional and behavioral disorders: A review of the literature from 1961 to 2000. *Journal of Emotional and Behavioral Disorders, 11*, 198-210.
- U.S. Department of Education. (2011-2012). *National Center for Education Statistics, Schools and Staffing Survey (SASS), "Public School Teacher Data File," 2011–12*. Retrieved from:
- Wagner, M. M. (1995). Outcomes for youths with serious emotional disturbance in secondary school and early adulthood. *Critical Issues for Children and Youths, 5*, 90– 112.
- 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 disturbances receiving special education. *Journal of Emotional and Behavioral Disorders, 13*(2), 79-96.
- Walker, B., Cheney D., Stage, S., Blum, C. & Horner, R. H. (2005). Schoolwide screening and positive behavior supports: Identifying and supporting students at risk for school failure. *Journal of Positive Behavior Interventions, 7*(4), 194-204.
- Walker, H. M., Nishioka, V. M., Zeller, R., Severson, H. H., and Feil, E. G. (2000). Causal factors and potential solutions for the persistent underidentification of students having emotional or behavioral disorders in the context of schooling. *Assessment for Effective Intervention, 26*(1), 29-29.
- Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices*. Belmont, CA: Wadsworth.
- Walker, H. M., Severson, H. H., & Feil, E. G. (2014). *Systematic Screening for Behavior Disorders* (2nd ed.) Eugene, OR: Pacific Northwest Publishing.
- Wallace, J. M., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991-2005. *Negro Educational Review, 59*(1-2), 47-62.

Wehby, J. H., Symons, F. M., Canale, J., & Go, F. (1998). Teaching practices in classrooms for students in emotional and behavioral disorders: Discrepancies between recommendations and observations. *Behavioral Disorders*, 24, 52-57.

APPENDIX
IRB Approval

ACTION ON PROTOCOL APPROVAL REQUEST



Institutional Review Board
Dr. Dennis Landin, Chair
130 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu | lsu.edu/irb

TO: Frank Gresham
Psychology

FROM: Dennis Landin
Chair, Institutional Review Board

DATE: March 2, 2016

RE: **IRB#** 3697

TITLE: Measuring Teachers' Perceptions of Student Behavior Using the Systematic Screening for Behavior Disorders

New Protocol/Modification/Continuation: New Protocol

Review type: Full ___ Expedited X **Review date:** 2/19/2016

Risk Factor: Minimal X Uncertain _____ Greater Than Minimal _____

Approved X **Disapproved** _____

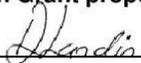
Approval Date: 2/29/2016 **Approval Expiration Date:** 2/28/2017

Re-review frequency: (annual unless otherwise stated)

Number of subjects approved: 16

LSU Proposal Number (if applicable):

Protocol Matches Scope of Work in Grant proposal: (if applicable) _____

By: Dennis Landin, Chairman 

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. **SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc.**

**All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at <http://www.lsu.edu/irb>*

APPENDIX

Administrator Consent Form



LOUISIANA STATE UNIVERSITY

ADMINISTRATOR INFORMED CONSENT FORM

We are requesting your approval and support to conduct the study *Measuring Teachers' Perceptions of Student Behavior Using the Systematic Screening for Behavior Disorders* at your school. The following sections outline the details of the study.

Purpose of the Study: As part of school- and district-wide preventative efforts, universal screening serves to identify students at risk for emotional and behavioral disorders while their behavior is still amenable to treatment. However, there are few universal screeners available for middle school students, who may be at heightened risk for developing emotional and behavioral difficulties due to major academic and social changes. This study is being conducted to examine a universal behavior screener recently adapted for use in middle schools. In addition, we will be investigating the comparative responses of teachers from different academic subjects.

Description of the Study: We are requesting approval to conduct a study on a universal behavior screener at your school. With your support, we will request participation in the study from teachers at your school. We will provide a training to teachers on the study procedures and provide all materials before the study begins. Students will *not* be observed, asked to complete anything or do anything differently, nor interact with researchers. Teachers will be asked to rank 10 of their students based on the severity of their problem behavior and complete two short measures of those students' behavior in school. In addition, teachers will provide an estimate the students' current academic grade and number of disciplinary referrals administered from their class. No identifying information, such as student names, will be collected. Completing the training, rating scales, and additional student information will take approximately 1.5 hours of each teacher's time over the course of 1-2 weeks. Frank Gresham, Ph.D., and Megan Erickson, B.A., of the Department of Psychology at Louisiana State University (LSU) are conducting this research.

Benefits: By participating in this study, your school will be contributing to the evaluation of universal behavior screeners and knowledge on the benefits and disadvantages of different screeners. Findings will be useful in providing insight into the development of screeners and implementation in middle schools. As a practical benefit to your school, completing the behavior screeners may help detect students in your school who are in need of extra supports. Your school will also gain knowledge on available universal behavior screeners for use at your school. In addition, to show our appreciation for teachers' time, effort, and assistance in our research efforts, we ask you if we may provide a "free dress" day to participating teachers. All participating teachers will also be entered into a raffle to

APPENDIX

Administrator Consent Form

win a \$50 gift card. In order to be eligible for this compensation, teachers must participate until the end of the study.

Risks: There are minimal risks associated with participation in this study. For example, teachers may feel uncomfortable rating students' behavior; however, they will be trained on rating procedures to minimize these risks. Additionally, by completing the screeners, teachers may identify students in need of extra behavioral supports. If pursued, this may require the use of school resources but for the benefit of the students.

Right to Refuse: Participation in this study is voluntary and your school will only be included if you agree to participate. You may choose to withdraw your school's participation at any time without affecting your relationship with your school or with LSU.

Privacy: Data will be kept completely confidential through the use of ID numbers, so that data cannot be linked to names. Results of this study may be published, but no names or identifying information will be included.

If you have any questions about this study, you may contact Dr. Frank Gresham at (225) 578-4663 or Megan Erickson at merick7@lsu.edu, Monday-Friday 8:00 a.m. - 4:30 p.m. If you have any questions about your child's rights or other concerns, please contact Dennis Landin, Chairman, Institutional Review Board, (225) 578-8692, irb@lsu.edu, www.lsu.edu/irb.

By signing this form, I acknowledge that I have read and understand the above information. I also acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

PLEASE CIRCLE ONE:

I give approval for teachers and students at my school to participate. **YES** **NO**

Name (please print): _____

Signature: _____ Date: _____

Phone Number: _____

Email: _____
(Please provide the email address you are most easily reached at.)

APPENDIX

Teacher Consent Form



LOUISIANA STATE UNIVERSITY

TEACHER INFORMED CONSENT FORM

We are requesting your participation and collaboration in the study *Measuring Teacher's Perceptions of Student Behavior Using the Systematic Screening for Behavior Disorders*. The following sections outline the details of the study.

Purpose of the Study: As part of school- and district-wide preventative efforts, universal screening serves to identify students at risk for emotional and behavioral disorders while their behavior is still amenable to treatment. However, there are few universal screeners available for middle school students, who may be at heightened risk for developing emotional and behavioral difficulties due to major academic and social changes. This study is being conducted to examine a universal behavior screener recently adapted for use in middle schools. In addition, we will be investigating the comparative responses of teachers from different academic subjects.

Description of the Study: We are requesting your assistance in this study on universal behavior screeners. We will provide a training on the study procedures and provide all materials before the study begins. You will be asked to rank 10 students in your class based on their behavior and for those students you will complete 2 measures of behavior and provide an estimate of the students' current academic grade and number of disciplinary referrals you have given from your classroom. Researchers will only be collecting student rank numbers for data analysis and not any identifiable information. On these measures, you will rate each student's behavior and indicate whether some the students have, to your knowledge, performed some particular behaviors. Completing the training, rating scales, and additional student information will take approximately 1.5 hours of your time over the course of 1-2 weeks. This study is being conducted with your administrator's approval. Frank Gresham, Ph.D., and Megan Erickson, B.A., of the Department of Psychology at Louisiana State University (LSU) are conducting this research.

Benefits: By participating in this study, you will be contributing to the evaluation of universal behavior screeners and knowledge on the benefits and disadvantages of a new screener for use in middle schools. Findings will be useful in providing insight into the development of screeners and implementation in schools. As a practical benefit to you, completing the behavior screeners may help you detect students in your classroom who are in need of extra supports, which should expedite and facilitate intervention planning and implementation. In addition, to show our appreciation for your time, effort, and assistance in our research efforts, we will provide each participating teacher with one "free

APPENDIX

Teacher Consent Form

dress" day sanctioned by your administrator. All participating teachers will also be entered into a raffle to win a \$50 gift card. In order to be eligible for this compensation, you must participate until the end of the study.

Risks: There are minimal risks associated with participation in this study. For example, you may feel uncomfortable rating students' behavior; however, you will be trained on rating procedures to minimize these risks. Furthermore, data will be kept completely confidential by associating your responses with an ID number.

Right to Refuse: Participation in this study is voluntary and you will only be included if you agree to participate. You may choose to withdraw your participation at any time without affecting your relationship with your school or with LSU.

Privacy: Data will remain completely confidential. Results of this study may be published, but no names or identifying information will be included.

If you have any questions about this study, you may contact Dr. Frank Gresham at (225) 578-4663 or Megan Erickson at merick7@lsu.edu, Monday-Friday 8:00 a.m. - 4:30 p.m. If you have any questions about your child's rights or other concerns, please contact Dennis Landin, Chairman, Institutional Review Board, (225) 578-8692, irb@lsu.edu, www.lsu.edu/irb.

By signing this form, I acknowledge that I have read and understand the above information. I also acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

PLEASE CIRCLE ONE:

I agree to participate. **YES** **NO**

Name (please print): _____

Signature: _____ Date: _____

Phone Number: _____

Email: _____
(Please provide the email address you are most easily reached at.)

VITA

A native of Spring, Texas, Megan Erickson received her bachelor's degree in psychology and English from Rice University in 2012. After graduating, she provided behavior therapy to children with autism and developmental disabilities for two years, before beginning the School Psychology Graduate Program at Louisiana State University in 2014. She anticipates receiving her master's in 2017, after which she intends to obtain certification as a Board Certified Behavior Analyst.