

5-13-2019

# An Examination of the Technical Adequacy, Classification Accuracy, and Usability of the SSIS SEL Screening/Process Monitoring Scales in Elementary School Populations

Shelby McCoy Byrd.

Louisiana State University and Agricultural and Mechanical College, shelby.byrd0514@gmail.com

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool\\_dissertations](https://digitalcommons.lsu.edu/gradschool_dissertations)

Part of the [School Psychology Commons](#)

## Recommended Citation

Byrd, Shelby McCoy ., "An Examination of the Technical Adequacy, Classification Accuracy, and Usability of the SSIS SEL Screening/Process Monitoring Scales in Elementary School Populations" (2019). *LSU Doctoral Dissertations*. 4912.  
[https://digitalcommons.lsu.edu/gradschool\\_dissertations/4912](https://digitalcommons.lsu.edu/gradschool_dissertations/4912)

This Dissertation 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 Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

**AN EXAMINATION OF THE TECHNICAL ADEQUACY, CLASSIFICATION  
ACCURACY, AND USABILITY OF THE *SSIS SEL SCREENING/PROGRESS  
MONITORING SCALES* IN ELEMENTARY SCHOOL POPULATIONS**

A Dissertation

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  
Doctor of Philosophy

in

The Department of Psychology

by

Shelby McCoy Byrd

B.S., Georgia College & State University, 2012

M.S., Valdosta State University, 2014

August 2019

## TABLE OF CONTENTS

ABSTRACT .....	iv
CHAPTER 1. REVIEW OF LITERATURE.....	1
Social Emotional Skills Defined .....	1
Emotional and Behavioral Problems .....	5
Social and Emotional Learning in Schools .....	8
Limitations of the Traditional Service Delivery Model .....	12
Paradigm Shift: Emphasis on Early Identification and Intervention.....	13
Universal Screening.....	16
Assessment of Social, Emotional, and Behavior Problems.....	24
Purpose and Research Questions.....	29
CHAPTER 2. METHOD .....	32
Participants and Setting .....	32
Measures .....	34
Procedure .....	40
CHAPTER 3. RESULTS.....	43
Descriptive Statistics .....	43
Reliability of the SSIS SEL Screening/Progress Monitoring Scales.....	45
Concurrent Validity .....	46
Conditional Probabilities and ROC Analyses .....	48
Predictive Validity .....	58
Social Validity .....	60
CHAPTER 4. DISCUSSION .....	62
Limitations and Future Directions.....	72
REFERENCES .....	75
APPENDIX A. TEACHER DEMOGRAPHIC FORM .....	81
APPENDIX B. STUDENT DEMOGRAPHIC FORM.....	82
APPENDIX C. SIBS AND SEBS COMBINED MEASURE.....	83
APPENDIX D. SSIS SEL EDITION – TEACHER FORM CORE SKILLS SCALE .....	84
APPENDIX E. USAGE RATING PROFILE – ASSESSMENT.....	85
APPENDIX F. STUDENT LEVEL DATA SHEET .....	86
APPENDIX G. SCHOOL ADMINISTRATOR INFORMED CONSENT FORM.....	87

APPENDIX H. TEACHER INFORMED CONSENT FORM .....	89
APPENDIX I. IRB APPROVAL .....	91
VITA.....	92

## ABSTRACT

Social emotional skills and competencies are integral to student success at home, school, and in the larger community. Extant research also consistently demonstrates that social emotional skill deficits are associated with various adverse outcomes. Universal screening for social emotional and behavioral risk in schools facilitates early identification and targeted intervention, with the primary goal to mitigate and reduce these potential adverse outcomes for students. Research on the technical adequacy and classification accuracy of universal screening is essential to this process to ensure efficient and accurate identification, as well as subsequent implementation of social emotional interventions targeting deficits in skills. The purpose of the current study was to extend existing research by Elliott et al. (in press), and further investigate the technical adequacy, classification accuracy, and usability of the SSIS SEL Screening/Progress Monitoring Scales. Results revealed evidence for the reliability, concurrent validity, and short-term predictive validity of the SSIS SEL Screening/Progress Monitoring Scales. Additionally, classification accuracy indices were adequate when compared to two well-researched criterion screening measures. Lastly, teachers rated the SSIS SEL Screening/Progress Monitoring Scales as largely acceptable, feasible and useable. In general, results extend the Elliot et al. (in press) findings and provide additional information on the predictive validity, classification accuracy, and usability of the SSIS SEL Screening/Progress Monitoring Scales in a sample of public elementary-school students and teachers. The following manuscript includes further examination of these results, a discussion of the importance of these findings, and implications for use in schools.

## CHAPTER 1 REVIEW OF LITERATURE

The United States educational system greatly emphasizes the development of academic competencies among elementary, middle, and high school students. Families, community members, and educators largely agree that schools should emphasize mastery in core academic subjects, as well as assist in the development of critical thinking and problem solving skills (Weissberg, Durlak, Domitrovich, & Gullotta, 2015). As demonstrated in extant research, social, emotional, and behavior problems significantly impact academic success and social functioning within schools (Gresham, 2005). The National Academy of Science estimates that while 60% students enter school with adequate cognitive skills, only approximately 40% of school-aged children enter kindergarten with the social and emotional skills needed to succeed in school. The emphasis on cognitive development and academic readiness among students has historically overshadowed the need for early emotional and social skill development (Raver, 2002). However, research investigating evidence-based interventions and the development of measures of social, emotional, and behavioral functioning has substantially increased.

### **Social Emotional Skills Defined**

***Social Skills.*** Social skills are defined as a particular class of behaviors that allow an individual to successfully complete a social task, such as facilitating communication with others, playing a game, or meeting new people (Gresham, 1986; Gresham & Elliot, 2014). Social skills or social emotional skills are terms often used interchangeably to illustrate an individual's ability to perform social behaviors appropriately, as well as manage and express emotions in socially and culturally acceptable ways. Social emotional competencies are integral to student success in and outside of school. Social skills enable students to navigate interpersonal domains at school, home, and in the larger community, and are involved in the development and maintenance of

interpersonal relationships and contribute to peer acceptance. Additionally, social skills are associated with positive school adjustment and the ability to cope and adapt in changing social environments (Gresham, Van & Cook, 2006).

***Social Competence.*** Although social skills contribute to social competence, no single social behavior is sufficient for competence. Social competence is an evaluative term based on another person's judgment of an individual's performance of a social task (Gresham, 1986). According to Gresham (1986), judgments of social competence might be based on the following three criteria: (1) relevant evaluation by peers, parents, or teachers; (2) evaluation of competence relative to pre-established criteria; and (3) evaluation of performance relative to a normative standard (e.g., scores on a standardized social skills measure). Judgments of social competence may be used to identify deficits and implement individualized instruction.

***Social Skill Deficits.*** Prior to implementing interventions for social skill deficits, it is important to consider the type of social skills deficit exhibited. Gresham (1981) was the first to delineate and describe two major types of social skill deficits, acquisition and performance deficits. This distinction has been further examined in research, and is widely accepted as a method of conceptualizing social skill deficits (Elliott & Gresham, 2014). An acquisition deficit is defined as lack of knowledge, inability to perform social behavior fluently, and difficulty determining which social skills to enact in different social settings. Acquisition deficits may result from an inability to discriminate between different social settings, social-cognitive deficits, and/or difficulties integrating behaviors fluently. Acquisition deficits can be characterized as being "can't do" problems, in which the student cannot perform the particular social behavior. In contrast, a performance deficit is described as failure to exhibit a particular social behavior, even though the student may know how to perform the social skill. Performance deficits are the result

of lack of motivation, rather than the lack of knowledge or acquisition problem. Performance deficits can be conceptualized as “won’t do” problems, in which the student knows how to do the behavior, but does not want to perform it (Elliott & Gresham, 2014; Gresham, 1981).

Discriminating between types of social skills deficits is important when implementing interventions. A student with acquisition deficits would likely benefit from explicit instruction of social skills. However, remediation of performance deficits would likely require manipulation of antecedents and consequences within the student’s social settings to maximize the probability of social skill acquisition and performance.

***Academic Enablers and Disablers.*** Social and emotional factors can facilitate or hinder student learning and success in school. Extant research has demonstrated significant and predictive relationships between student social emotional behavior and academic achievement (Diperna & Elliot, 2002; Gresham & Elliot, 2014). Social skills create an academic climate that is conducive to learning, with positive peer interactions and social behaviors being associated with greater academic engagement and higher levels of achievement (Wentzel, 1991; Wentzel, 2009). Malecki and Elliott (2002) indicated that social skills are positively predictive of concurrent academic achievement in elementary students, whereas problem behaviors are negatively predictive of concurrent academic achievement in elementary students. Social skills were also significantly predictive of future academic performance (Malecki & Elliot, 2002). Additionally, Caprara and colleagues concluded that teacher-rated social behavior in third grade was a better predictor of future academic success (i.e., eighth grade academic achievement) than third grade academic achievement (Caprara, Barbaranell, Pastorelli, Bandura, & Zimbardo, 2000).

Rabiner, Godwin, and Dodge (2016) extended previous research by examining the relationship among social competence, attention, and academic achievement across student



academic career (i.e., academic achievement at end of elementary school and academic grades during sixth, seventh, and eighth grade) and into young adulthood (i.e., age 24 or 25). In this particular study, social competence was operationalized and computed using peer social acceptance ratings and peer-rated aggression. Low social preference was predictive of lower academic performance at the end of elementary school and through middle school. Lower social preference scores were also predictive of fewer years of education attained at young adulthood (Rabiner, Godwin, & Dodge, 2016). Reported social skills and behavioral problems have also been used to determine student readiness for school, with greater social skills being associated with promotion to first grade (Agostin & Bain, 1997).

For these aforementioned reasons, Gresham and Elliott (2014) described social skills as “academic enablers” that allow students to engage with and benefit from academic instruction within the classroom. In contrast, social emotional deficits disrupt the classroom environment for all students and are associated with lower academic performance (Benson, 2006). Students who fail to develop adequate social competencies are at risk for a host of negative outcomes, including increased aggressive behaviors, peer rejection, loneliness, social dissatisfaction, poor academic performance, school dropout, substance abuse, difficulties maintaining employment and relationships, and criminality (Maar, 2006; Rutherford, et al., 2004). Whereas social emotional skills function as academic enablers, competing problem behaviors function as “academic disablers” (Gresham & Elliott, 2014). This is particularly true for externalizing behavior problems, such as aggression, noncompliance, and teacher defiance. In contrast to social emotional skills, externalizing problem behaviors are associated with decreased academic engagement and achievement. Although research has not consistently determined if academic deficits are correlates (i.e., moderators), causes (i.e., mediators), or consequences of

externalizing problem behaviors, researchers agree that problem behaviors greatly exacerbate academic difficulties (Gresham & Elliot, 2014).

### **Emotional and Behavioral Problems**

Social and emotional behaviors range on a continuum from deficits to competencies, with suggested emotional and behavioral problems occurring in approximately 1 in 5 students (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; National Association of School Psychologists [NASP], 2002). Students with emotional and behavioral disorders (EBDs) exhibit various behavioral patterns that adversely affect their interpersonal and/or academic functioning. Researchers largely concentrate on two broad dimensions of problem behaviors exhibited by those with EBDs, externalizing and internalizing behaviors. Externalizing behaviors are overt behaviors that are directed outward and toward the environment or other people. Externalizing behaviors are often conceptualized as undercontrolled behaviors (Cook, Volpe, & Gresham, 2012; Wiley & Siperstein, 2015). Examples of externalizing behaviors include verbal and physical aggression, disruptive behavior (e.g., leaving seat without permission, blurting out), tantrums, noncompliance, defiance, and destruction of property. Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD) are DSM-5 diagnoses that are commonly associated with externalizing patterns of behaviors (American Psychiatric Association [APA], 2013). In contrast, internalizing behaviors are covert behaviors that are directed internally or toward self. Internalizing behavior patterns are conceptualized as overcontrolled behaviors (Cook et al., 2012; Wiley & Siperstein, 2015). Examples of internalizing behavior patterns include social withdrawal, sadness, and worries. Generalized Anxiety Disorder (GAD), Major Depressive Disorder (MDD), and Social Anxiety Disorder

(SAD) are DSM-5 diagnoses that are commonly associated with internalizing patterns of behaviors (APA, 2013).

The Individuals with Disabilities Education Improvement Act (IDEIA) of 2004 defines 13 special education categories that students may be classified. Students with emotional and behavioral disorders qualify for special education services with the classification of Emotional Disturbance (ED). According to the federal definition, emotional disturbance is a condition exhibiting one or more of the following characteristics over an extended period of time and to a marked degree, which adversely affects educational performance: (a) an inability to learn that cannot be explained by intellectual, sensory, or health factors; (b) an inability to build or maintain satisfactory interpersonal relationships with peers and teachers; (c) inappropriate types of behavior or feelings under normal circumstances; (d) a general pervasive mood of unhappiness or depression; or (e) a tendency to develop physical symptoms or fears associated with personal or school problems (Individuals with Disabilities Education Improvement Act [IDEIA], 2004). Unlike other disabilities outlined in IDEIA, impairment in social and emotional functioning is central to the definition of emotional disturbance. Social and emotional problems are directly linked to problem behaviors in students classified as ED (Wiley & Siperstein, 2015). According to the U.S. Department of Education, only approximately 1% of school-aged children and adolescents receive services under the ED classification (Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005). However, this is significantly less than the estimated 20% of school-aged children and adolescents with mental health problems (Gresham, 2005).

***Outcomes Associated with Emotional and Behavioral Problems.*** Extant research examining developmental trajectories of emotional and behavioral problems from childhood through adolescence and young adulthood largely indicates negative short-term and long-term

outcomes that affect the student, classroom, school, and larger community. The characteristics, functioning, and challenges of students with EBDs have been examined in two longitudinal studies, the Special Education Elementary Longitudinal Study (SEELS) and the National Longitudinal Transition Study-2 (NLTS2). The SEELS was conducted with elementary and middle school students, and the NLTS2 was conducted with high school students and young adults. In summarizing these two milestone longitudinal studies, Wagner et al. (2005) reported that students with ED were significantly more likely to exhibit social skills and communication deficits compared to peers with other disabilities served under IDEIA. Additionally, children with EBDs exhibit more self-regulation deficits than same-aged peers, including deficits in planning, directing, and controlling emotions, thoughts, and behaviors (Barkley, 2010). Relatedly, children with EBDs are less likely to exhibit social and emotional competencies, including interpersonal problem solving, emotion expression, and social and relationship skills (Graziano & Hart, 2016).

Bradley, Doolittle, and Bartolotta (2008) indicated that students with EBDs exhibit greater academic disengagement and poor work completion when compared to peers. Additionally, students with EBDs exhibit high rates of behaviors resulting in school discipline, including office discipline referrals (ODRs), suspensions, and expulsions. School absences, academic failure, school dropout, and lack of postsecondary education attainment have also been found to be associated with EBDs (Bradley et al., 2008; Wagner et al., 2005). Using data from the Children of the National Longitudinal Surveys of Youth data set, McLeod and Fettes (2007) also demonstrated that internalizing and externalizing problems in childhood and adolescence were significantly associated with poor educational attainment (i.e., high school graduation and college enrollment). The various adverse academic and social outcomes associated with EBDs

explicate the importance of early identification of risk and implementation of intervention services.

### **Social and Emotional Learning in Schools**

***Social Emotional Learning.*** Social emotional learning (SEL) has been defined as the “process of acquiring knowledge, skills, attitudes and beliefs to identify and manage emotions; to care about others; to make good decisions; to behave ethically and responsibly; to develop positive relationships and to avoid negative behaviors” (Elias & Mocerri, 2016, p. 424). SEL is process of learning how to experience, express, and regulate positive and negative emotions in socially acceptable ways (Ashdown, 2012). The term “social and emotional learning” was first introduced at the Fetzer Group meeting in 1994. At this meeting, educators, researchers and advocates developed an international organization to assist in the establishment and integration of SEL programs within schools—The Collaborative for Academic, Social, and Emotional Learning (CASEL). Since that time, CASEL has worked to emphasize the need for social-emotional development within universal education setting, and components of SEL programs and curriculum have been widely researched. CASEL’s mission is to provide evidence-based explicit instruction of social skills to preschool, elementary, middle and high school students. Additionally, it attempts to advance SEL research, translate and disseminate scientific knowledge, enhance training, and increase collaboration among educators, researchers, policymakers, and advocates to expand SEL efforts (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2016; Payton et al., 2000).

CASEL has further refined and simplified the definition of SEL to five core competence domains, including self-awareness, self-management, social-awareness, relationship skills, and responsible decision making. Competence in self-awareness involves the ability to recognize

one's emotions and thoughts accurately, as well as understanding how emotions and thoughts influence behaviors. Self-awareness also includes an understanding of one's personal goals, values, and strengths and weaknesses. Additionally, self-awareness should be grounded in optimism and self-efficacy. Self-management is the ability to regulate one's thoughts, emotions, and behaviors effectively in different situations and settings. Competence in the self-management domain requires that the individual be able to delay gratification, manage stress, regulate impulses, and persist through challenging situations in order to achieve personal and educational goals. Social awareness is the ability to take the perspective of and feel compassion and empathy toward other persons of diverse cultures and backgrounds. Competence in social awareness requires knowledge of social norms and practices, as well as recognition of family, educational, and community resources and supports. Relationship skills include the ability to develop and maintain interpersonal relationships. Competence within this domain requires that the individual be able to effectively communicate, listen to others, cooperate, resist inappropriate social pressures, resolve conflicts, and seek assistances when needed. Responsible decision-making involves the capacity to make constructive decisions about personal and social behaviors based on ethical responsibility, safety issues, and social norms. To make decisions responsibly, the individual must identify the problem, analyze the situation, and develop hypotheses and solutions to solve the problem. Competence within this domain requires that the individual realistically evaluate consequences of his or her behaviors, and consider the well-being of self and others prior to and when making decisions. Together, these five competencies increase an individual's ability to interact positively with various people in different contexts.

CASEL purposely included “learning” in the term “social and emotional learning” to emphasize that the acquisition of core social skills and competencies (i.e., self-awareness, self-

management, social awareness, relationship skills, and responsible decision making) is a learning process, and one that can be easily integrated within the core curriculum at schools (Weissberg, Durlak, Domitrovich, & Gullotta, 2015). SEL programs should be evidence-based and integrated within regular classroom instruction. Durlak et al. (2010, 2011) purports that successful SEL programs promote social and emotional competences through four elements illustrated by the acronym SAFE: sequenced activities, active participation, focused instruction, and explicit teaching of skills. Specifically, SEL approaches should include *sequenced* and connected learning activities to promote skill development; integrate *active* forms of learning; emphasize *focused* instruction of personal and social skills; and *explicitly* target social and emotional skills. Consistent with other emotional and behavioral problems, early intervention is key. Effective SEL programming should begin in preschool and continue through elementary, middle, and high school (Weissberg & Cascarino, 2013). Furthermore, Weissberg et al. (2015) suggest that social, emotional, and academic development is further enhanced through coordinated classroom, school, family, and community efforts.

***Outcomes of SEL Programs.*** Schools play an important role in equipping children and adolescents with the abilities to effectively communicate and interact with others in socially and emotionally skilled ways. Applied research has demonstrated the positive outcomes and effects of SEL implementation in schools for a variety of student populations. The potential benefits of improving student behavior and well-being are vast and long term. Short-term proximal effects of SEL implementation in schools include improved academic performance, lower grade retention, greater class engagement, and increased motivation (Jones, Greenberg, Crowley, 2015). Ashdown (2012) demonstrated that teacher facilitated SEL curriculum (i.e., the *You Can Do It! Early Childhood Education Program*) resulted in significant reduction in total problem

behaviors exhibited by first graders (i.e., externalizing, internalizing, and hyperactivity), as well as gains in positive emotions, social skills, and social-emotional competence of students.

Additionally, results indicated a significant increase in overall reading achievement among the students that participated (Ashdown, 2012). McBride, Chung, and Robertson (2016) examined the effects of school-based social and emotional learning in middle school students. Results indicated a significant reduction in failing grades and skipping class among the participating students, extending research on positive effects of SEL implementation to middle school students (McBride, Chung, & Robertson, 2016).

A meta-analysis of 213 studies evaluating SEL programs further illustrated the benefits of applying SEL curricula in elementary, middle, and high schools in both urban and rural schools in the United States (Durlak, Weissberg, Dymicki, Taylor, & Schellinger, 2011). Approximately half (47%) of the reviewed interventions utilized randomization to assess effects. Primary outcomes of SEL programming included increases in social and emotional skills of students, development of prosocial behaviors or attitudes, and greater mental health. Further, implementation of SEL was associated with greater academic performance, including an 11-percent-point increase in achievement scores as reflected through report card and test scores (Durlak et al., 2011). Another meta-analysis of 75 studies examining after school programs that foster personal and social skills further explicated the impact of applying SEL programs with school-aged children (Durlak, Weissberg, & Pachan, 2010). Outcomes associated with after-school SEL programs include greater self-perception, increased bonding and connectedness to school, development of positive social behaviors, reduction in problem behaviors, and increased academic achievement.



Researchers have also postulated that positive effects of SEL implementation extend beyond improvement in academic scores and student classroom behaviors. Jones, Greenberg, and Crowley (2015) emphasized the potential economic benefit for SEL programming in schools, as possible long-term distal effects of school SEL interventions include lower mental health costs, employment success, and increased quality of life. Additionally, Greenberg, Katz, and Kelin (2015) purported that SEL program benefits might even extend to overall health of recipients. Specifically, researchers posited that self-regulation skill development might change how children interact with the environment, reducing long-term effects of stress and creating a lasting effect on biological systems (Greenberg et al., 2015).

### **Limitations of the Traditional Service Delivery Model**

Prior to the reauthorization of IDEIA in 2004, schools operated under the traditional “wait to fail” model of identification and intervention for students with academic, behavioral, or social difficulties (Walker, Severson, & Seeley, 2010). The traditional identification and service delivery model relied on teacher nominations, which often resulted in students not being referred for special education services until behaviors were no longer manageable within the regular education classroom (Gerber & Semmel, 1984). Additionally, once student problem behaviors peaked unmanageable levels, teachers would often employ the assistance of other school personnel (e.g., behavior specialists, school guidance counselors, school psychologists) to assume responsibility for addressing and managing student behaviors inside and outside of the classroom. This sequence of events often concluded in a “refer-test-place” decision that removed the disruptive student from the regular education classroom. The “wait to fail” model too often allowed for minor problem behaviors to develop into severe problem behaviors prior to identification and implementation of services, and thus resulted in loss of critical opportunities

for early intervention and put students with emotional and behavioral problems at risk for many short-term and long-term negative outcomes (Walker et al., 2010).

Although teachers are likely to have expansive knowledge of student behaviors in their classrooms, there are several limitations associated with relying solely on teacher nomination for referral for additional services, primarily related to inconsistent and inaccurate identification of students (Walker et al., 2010). Teacher referrals are often highly idiosyncratic and rely heavily on subjective judgment, thus contributing to inaccurate and inconsistent referrals within the classroom. Additionally, students with emotional and behavioral disorders may exhibit a range of behaviors, from internalizing to externalizing problem behaviors. Because externalizing problem behaviors, such as disruption and noncompliance, are much more salient in the classroom, teachers are more likely to refer students with externalizing behaviors patterns (Dowdy, Doane, Eklund, & Dever, 2011). In contrast, teachers may be unaware of student internalizing problems (e.g., anxiety, depression) because of the lack of associated classroom disruption. Therefore, students with internalizing patterns of behavior are often overlooked and thus are not referred for additional services (Walker et al., 2010).

### **Paradigm Shift: Emphasis on Early Identification and Intervention**

Extant research demonstrates that early identification of social, emotional, and behavioral problems is key to improving outcomes for at-risk students. Systematic universal screening and early identification of at-risk students has become widely acknowledged as an important professional practice within schools (Walker et al., 2010). The shift toward embracing prevention and early identification efforts in schools was partly initiated with federal policy demanding educational reform. Since the National Commission on Excellence in Education (1983) published *A Nation At Risk*, individuals in the educational community have advocated for

the delivery of high-quality instruction to all students through the use of evidence-based practices. The United States Congress further endorsed this initiative with the passing of the No Child Left Behind Act (NCLB) in 2001, which encouraged schools and districts to adopt and implement evidence-based programs (Stoiber, 2014). Additionally, IDEIA specified the importance of early identification of academic and behavioral problems for the purpose of reducing the need for special education services at a later time (Stoiber, 2014; Walker et al., 2010). IDEIA purports that schools have an obligation to establish systems that accurately identify students in need for additional services. IDEIA also indicates that schools may allocate up to 15% of available funds for the use of early screening, intervention, and prevention services (Walker et al., 2010).

***Multitiered Systems of Support.*** Schools have largely recognized the need for multitiered service delivery models as an alternative to the traditional “wait to fail” model. Multitiered systems of support (MTSS) is broadly defined as a multicomponent, integrated, and comprehensive positive support system through which students with academic or behavioral problems are identified and provided with evidence-based instruction and support (Stoiber, 2014). The primary goal of MTSS is to prevent, reverse, and reduce mental health problems, while simultaneously fostering social, emotional and academic success among all students in a school (Strein, Hoagwood, & Cohen, 2003).

Parallel to the public health prevention model, MTSS in schools employs a tiered method of instructional delivery, with a three-tiered system being most common. The three tiers are typically called (1) primary or universal, (2) secondary or targeted, and (3) tertiary or intensive. The three tiers within MTSS are differentiated based on intervention, with greater and more intensive support being provided as students move up the three tiers. High-quality screening

should be implemented at the universal level to determine social emotional needs of all students in the school. Additionally, within the MTSS framework, universal supports (e.g., social-emotional learning programs) are delivered to all students to promote academic, social, and emotional success. Tiers 2 and 3 are not intended to replace universal instruction or support, but to further enhance and supplement student learning (Stoiber, 2014). Screening and brief assessment tools are also useful in secondary and tertiary tiers to determine student responsiveness to interventions.

***SEL and PBIS.*** Under the umbrella of MTSS, social emotional learning and positive behavior interventions and supports (PBIS) are two widely researched approaches to target student behavior in schools. Both SEL and PBIS emphasize system changes and focus efforts on prevention. As stated previously, SEL programs teach skills related to self-management, interpersonal problem solving, emotional regulation, and social awareness. Under a three-tiered approach to alleviate social-emotional problems, Tier 1 focuses broadly on prevention and early identification, and includes both universal screening for social-emotional competence and the implementation of classwide social-emotional learning instruction. Tier 2 represents more focused and targeted interventions, such as small group instructional strategies to promote self-regulation skills. Tier 3 services are the most intensive and individualized, and may include individual counseling or one-on-one behavioral consultation within the home. These tertiary services are provided to students who demonstrate significant and persistent signs of social emotional deficits and behavioral problems (Squires, 2010).

PBIS is a practical approach to target inappropriate behaviors by addressing contextual factors surrounding the problem behavior (Stoiber, 2014). PBIS is grounded in applied behavior analysis and utilizes instruction, modeling, and reinforcement of appropriate behaviors and

implementation of behavioral management systems to decrease inappropriate behaviors (Cook et al., 2015). The overall goal of PBIS at the universal level is to teach and support appropriate student behaviors, and prevent initial instances of inappropriate behaviors. Major strategies at Tier 1 include establishing and defining behavioral expectations for all students, modification of the environment to prevent or reduce instances of problematic behaviors (e.g., increasing proximity, active supervision), and applying reinforcement of appropriate behaviors and punishment-based consequences for inappropriate behaviors. Consistent with other multitiered models of support, PBIS utilizes screening and continuous progress monitoring data for decision making. Tier 2 and 3 supports increase in intensity for those that universal supports are not adequate (Frey, Lingo, & Nelson, 2010).

Researchers have begun to examine the impact and acceptability of integrated evidence-based practices on student outcomes, including SEL and PBIS programs. Cook and colleagues (2015) examined the effect of SEL and PBIS integration on student mental health. Results indicated that a combination of both PBIS and SEL produced greater improvements in reported mental health and decreases in externalizing behaviors when compared to SEL alone, PBIS alone, and control conditions. Cook et al. (2015) findings provide further evidence for a comprehensive, integrated system of support to address student problem behaviors and foster appropriate social-emotional skills and behaviors.

### **Universal Screening.**

Universal screening is an alternative approach to the remediation-based traditional model, and can be easily incorporated into the current preventative, multitiered service delivery framework within schools (Dowdy et al., 2015). Universal screening in schools is broadly defined as the administration of measures or collection of relevant data to provide estimations of

current and future performance at the individual and group level (e.g., classroom, school, district). Universal screeners are designed to (1) be administered to all students within the school; (2) identify students at risk for academic, behavioral, or emotional problems and thus in need of preventative or intervention services (e.g., social-emotional learning, small group academic instruction); (3) provide data to determine if school-wide programs and initiatives (e.g., social-emotional programs, positive behavioral supports, core academic instruction) are effectively meeting the needs of all or the majority of students; and (4) provide information regarding the individual student academic, behavioral, and social-emotional needs (Albers & Kettler, 2014). From the preventative perspective, a major goal of universal screenings is to identify students who are currently asymptomatic, but will likely experience behavioral, emotional, or academic difficulties in the future (Albers & Kettler, 2014). Therefore, the adoption of universal proactive screening procedures is likely to address the problems associated with reactive referral methods, including the underreferral of internalizing behavior problems and the inaccuracy of teacher referrals (Walker et al., 2010).

***Approaches to Universal Screening.*** The educational system has employed screening methods to some degree for many years. However, the major difference between the current and previous methods is the present emphasis on universal and standardized methods of assessment. Traditionally, “screening” approaches tended to consist of (1) records reviews; (2) behavioral observations; (3) in-depth report from students, teachers, and parents via rating scales and/or interviews; and (4) overall academic performance on school or district level assessments. As mentioned previously, many of these methods are flawed and often resulted in misidentification of students. Current best practices suggest that school administrators, educators, and psychologists utilize standardized procedures of screening that are reliable and valid. Such

methods might include broadband, narrowband, or multi-gate approaches (Albers & Kettle, 2014).

Broadband assessments are designed to evaluate several domains concurrently, such as academic issues, externalizing behaviors, and internalizing behaviors. A commonly used broadband measure in schools and clinics is the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla, 2001). Narrowband approaches are constructed to assess a specific domain of interest, such as anxiety or oppositional behavior. The State-Trait Anxiety Inventory for Children (STAI-C; Spielberger, Edwards, Lushene, Montuori, & Platzek, 1973) and the Multidimensional Anxiety Scale for Children (MASC; March, 1997) are examples of narrowband assessment used to assess anxiety in children and adolescents. Broadband assessments are useful to capture a broad representation of a student's current functioning, whereas narrowband assessments provide a more in-depth assessment of a specific target area (Albers & Kettler, 2014).

Schools may also adopt multi-gating approaches to screening and assessment. According to Albers and Kettler (2014), all students are initially screened with a brief measure within the multi-gate process. Sequential stages are then completed, and more intensive assessments are administered to students identified as at-risk in previous stages of the assessment. In the initial stage of measurement (i.e., Gate 1), the classroom teacher completes a ranking measure of all students in his or her classroom according to the frequency of a problematic behavior (e.g., disruptive behaviors). Next, the teacher completes a standardized, broadband behavior rating on students identified as at-risk during the initial phase of the assessment (i.e., Gate 2). Lastly, for the students identified in the second stage, the school psychologist completes systematic direct observations and administers additional rating scales to the parent and student (i.e., Gate 3). The

multi-gating method is advantageous because it employs multi-methods (e.g., screeners, broadband behavior rating scales) and different raters (e.g., teachers, parents, students) in the assessment and identification process. Additionally, the multi-gate method is time and cost efficient, and increases the accuracy of identification of students in need of additional services (Alber & Kettlers, 2014).

***Considerations: Technical Adequacy and Usability of Screening Measures.*** The American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) released explicit guidelines or criteria for evaluating the technical adequacy of assessment instruments in the *Standards for Educational and Psychological Testing* (AERA et al, 1999). Glovers and Albers (2007) reviewed these criteria as they relate to universal screening in schools. Researchers purport that schools should evaluate the appropriateness of the screener intended for use (Glover & Albers, 2007). That is, the screener should be highly compatibility with school needs, align with relevant constructs of interests, have theoretical and/or empirical support of format and content, and possess a general fit with population group of interest (Glover & Albers, 2007). However, although a screening measure may be deemed as appropriate for intended use in a particular school, it is not useful if it cannot reliably and accurately predict the behaviors of the population group of interest. Therefore, schools should also evaluate the technical adequacy of screeners, including the adequacy of the normative sample demographics, reliability, and validity evidence (Albers & Kettler, 2014; Glover & Albers, 2007). Additionally, it is important that schools utilize universal screeners that are generally rated as highly usable, feasible, and acceptable among raters (Glover & Albers, 2007).



It is important to consider the adequacy of a normative sample when making inferences about a student's risk status. Screening instruments may provide local (e.g., school, district, or state) or national norms. Although national norms are typically more stable, local norms may be used to provide a better representation of the target students within their district and may lead to more meaningful implications for implementation of intervention or related services (Glovers & Albers, 2007). Glovers and Albers (2007) stated that when evaluating screening instruments, schools psychologists and school administrators should assess the adequacy of the sample representativeness, recency, and sample size. Specifically, the normative sample should be representative of the target population, such as in age, gender, and race/ethnicity. The normative sample size should be large enough to include students of varying age and grade or performance levels. Additionally, the recency of norms should be considered when using screening instruments for identification in schools (Glover & Albers, 2007).

When implementing universal screening systems in schools, school psychologists or administrators should examine indices of reliability to determine consistency of the screening scores across the measure, time, and raters. Three types of reliability evidence often considered in research when evaluating screeners include internal consistency, test-retest reliability, and interrater reliability (Albers & Kettler, 2014; Glover & Albers, 2007). Internal consistency reliability is a measure of the relationship or correlation among different items on the same assessment instrument, and demonstrates if items or groups of items measure the same construct. Chronbach's coefficient alpha is a measure of internal consistency (AERA et al., 1999). Test-retest reliability measures the consistency or stability of results over time. Test-retest estimates are obtained by administering the assessment instrument to the same rater at two different points in time (AERA et al., 1999) Interrater reliability measures agreement among two or more raters,

and demonstrates how consistently different informants rate the same behavior (AERA et al., 1999). Although comprehensive assessment batteries should obtain a reliability coefficient of approximately 0.9, lower coefficient alphas are generally considered acceptable for screener systems because of the brief, low-stakes, and preventive focus. Specifically, reliability coefficients of approximately 0.7 or 0.8 are considered acceptable for screeners (Albers & Kettler, 2014). Additionally, because a screening measure cannot be valid if it is not reliable, researchers purport that reliability estimates should be examined prior to establishing validity (Albers & Kettler, 2014).

In addition to producing consistent or reliable results, a screening tool must also be accurate. Validity refers to the degree to which a measure assesses what it is intended to assess. The *Standards for Educational and Psychological Testing* outlined five types of validity evidence that should be considered when evaluating assessment procedures, including content validity, validity based on response processes, internal structure validity, validity based on relationships to other performance variables, and consequential validity (AERA et al., 1999). However, content validity, construct validity, and criterion-related validity are the three most commonly used types of validity evidence when evaluating screeners and other methods of assessment (Glover & Albers, 2007). Content validity is defined as the degree to which the content of the measure (i.e., assessment items) is appropriate and adequately reflects what it is intended to measure. Construct validity refers to the degree that the assessment instrument measures the construct that it is designed to measure (e.g., social skills, depressive symptoms). Two types of construct validity are convergent validity (i.e., positive relationship with similar variables) and discriminant validity (i.e., negative relationships with dissimilar variables). Lastly, criterion-related validity indicates the degree to which scores on a particular assessment predict

performance on a related criterion variable or outcome. Predictive and concurrent validity are two subtypes of criterion-related validity (Albers & Kettler, 2014; Glover & Albers, 2007).

Criterion-related validity is often used in academic and psychological assessment research because the relationship between screener performance and outcome (e.g., special education classification, academic failure, school discipline referrals) demonstrates the importance of early identification of students at risk for adverse outcomes. According to Glovers and Albers (2007) predictive validity is possibly the most important indicator of screener adequacy. Specifically, for a screening measure to be effective, it must be able to accurately discriminate between students who will and students who will not have academic or behavioral problems (Glover & Albers, 2007). Additionally, many school psychologists and educators rely on conditional probability indices to evaluate screening systems, which provide indicators of diagnostic or classification accuracy of a screener. Conditional probability indicators are proportions that illustrate two dichotomies: (1) the student may or may not be identified as needing intervention and (2) the student may or may not actually need intervention (Albers & Kettler, 2014). The two dichotomies could be further sorted into four categories within a screening outcome matrix: true positive (i.e., identified, actual need), false positive (i.e., identified, no need), false negative (i.e., not identified, actual need), and true negative (i.e., not identified, no need) (Albers & Kettler, 2014; Glover & Albers, 2007).

Based on the above four category outcomes, various indicators of conditional probability may be calculated, including the sensitivity, specificity, positive predictive value, and the negative predictive value. The sensitivity index demonstrates the proportion of individuals in need of intervention services that are accurately identified (Albers & Kettler, 2014; Glover & Albers, 2007). In contrast, the specificity index is an indicator of the proportion of individuals

that do not need services that are accurately not identified. The positive predictive value is the proportion of individuals accurately identified as at-risk (i.e., true positive), and the negative predictive value is the proportion of individuals accurately not identified that do not need help (i.e., true negative) (Albers & Kettler, 2014; Glovers & Albers, 2007). Additionally, educators and school psychologists may assess hit rates (i.e., proportion of true positives and true negatives of total students) and base rates (i.e., proportion of all students that actually need intervention of total number students assessed) to provide further context for interpreting different indices of conditional probabilities (Albers & Kettler, 2014).

Though a screener may be technically sound, it is not likely to be consistently used in schools unless it is perceived as practical or usable within the context. Glovers and Albers (2007) outlined six considerations for evaluating usability of a screening measurement. First, the cost of screening implementation must not outweigh the benefits. Use of a screener should not place an unreasonable human or financial burden on the school or users. Secondly, administration of a screening instrument should be feasible (Glovers & Albers, 2007). Often times, individuals that complete a universal screening instrument range in qualification and training backgrounds, and therefore instructions should be clear to all users. Complicated scoring and interpretation should be comprehensively explained, and the screener format should be suitable for the setting and target population. Third, the screening measure should be rated acceptable to variety of stakeholders, including school staff, administrators, teachers, and parents (Glovers & Albers, 2007). All users must agree that the benefits of screener usage outweigh the costs (e.g., time, financial cost, and stress of administration). Glovers and Albers (2007) assert that obtaining “buy in” from stakeholders is likely to increase the probability that a screening instrument is

consistently used over time to guide decision-making process and service delivery of intervention in schools.

Forth, it is also important to consider the required infrastructure for collecting, organizing, and interpreting assessment data in schools (Glovers & Albers, 2007). Though school staff may be available to facilitate the screening process, often personnel are required to take on new roles. As such, additional school staff may be needed to facilitate the universal screening implementation process. Fifth, consideration must also be given to whether or not appropriate accommodations are available for the targeted population (Glovers & Albers, 2007). For example, if target population includes English Language Learners, schools should ensure that adequate administration, scoring, and interpretation instructions are provided (AERA et al., 1999). Lastly, and possibly most importantly, information obtained from screening instruments should be useful to schools and result in improved treatment utility. Specifically, intervention recommendations generated from screening results should also be feasible, relevant, and acceptable within the school setting (Glovers & Albers, 2007).

### **Assessment of Social, Emotional, and Behavior Problems.**

Since the adoption of MTSS in schools, universal screening for early identification of emotional, behavior, and academic risks has become increasingly more widespread (Kamphaus et al. 2014). There are a variety of available screening measures for social, emotional, and behavioral problems. Some of the more prominent and heavily researched screening measures include: BASC-2 Behavioral and Emotional Screening System (BESS; Kamphaus & Reynolds, 2007), Social Skills Improvement System – Performance Screening Guide (SSIS-PSG; Elliott & Gresham, 2007), Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1992), Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001), and the Student Risk

Screening Scale (SRSS; Drummond, 1994). Additionally, recent studies further suggest technical adequacy and utility of the Student Internalizing Behavior Screener (SIBS; Cook et al., 2007) and Student Externalizing Behavior Screener (SEBS; Cook et al. 2012) in screening for emotional and behavioral concerns (Hartman, Gresham, & Byrd, in press).

The Social Skills Improvement System – Performance Screening Guide (SSIS-PSG) is a universal screening instrument that measures behavioral and academic indicators of school performance (Elliott & Gresham, 2007). The SSIS-PSG is the class-wide universal screening component of the Social Skills Improvement System (SSIS), which allows social skills and behaviors to be contextualized along with academic skills within a multitiered model. The SSIS-PSG is completed by teachers and can be used for screening with preschool, elementary, and secondary students. Skill domains assessed on the SSIS-PSG include Prosocial Behavior, Motivation to Learn, Reading Skills, and Math Skills. To complete the SSIS-PSG, teachers assign a single rating for each skill domain from 1-5 based on each student’s current level of functioning. Descriptions of performance levels and defining behaviors in each domain are provided to anchor ratings. Ratings correspond to risk levels as follows: 4 or 5 indicates expectations are met or exceeded based on age and grade level, 2 or 3 indicates moderate concern or difficulty, and a 1 suggests high level concern or risk. Ratings may also be linked to intervention.

Standardization studies completed by Elliott and Gresham (2007) included data from elementary to high school teachers and students. In a subset of elementary teachers and students, strong to moderate evidence of reliability was demonstrated with test-retest reliability coefficients ranging from  $r = .68$  to  $r = .74$  and interrater reliability coefficients falling between  $r = .55$  and  $r = .68$ . Additionally, psychometric evidence for concurrent validity of the SSIS-PSG

with the SSIS-Rating Scales (SSIS-RS) was demonstrated by a significant negative correlation between SSIS-PSG ratings and the SSIS-RS Problem Behaviors subscale and by significant positive correlations between SSIS-PSG Reading Skills, Math Skills, and Motivation to Learn ratings and SSIS-RS Social Skills subscale scores. Teachers rated the behaviors rated by the SSIS-PSG as important. The SSIS-PSG was rated as easy to understand and complete (Elliott & Gresham, 2007).

Miller and colleagues (2015) provided additional psychometric and usability evidence of the SSIS-PSG as compared to other measures of social emotional functioning, including the Direct Behavior Rating-Single Item Scales (DBR-SIS), Behavioral and Emotional Screening System (BESS), office discipline referrals (ODRs), and school nomination methods.

Approximately 2,000 students were assessed tri-annually (i.e., fall, winter, spring) by their respective teachers. Results indicated that teacher ratings using the DBR-SIS, BESS, and SSIS-PSG related in greater proportion of identified at-risk students than school nomination or ODRs. The SSIS-PSG Prosocial Behavior and Motivation to Learn subscales accurately identified 83% to 91% of students at-risk for social, emotional, or behavioral program during the academic year using the BESS as a criterion. Additionally, strong correlations were found between all rating scales, with the strongest correlations being between the BESS and SSIS-PSG ranging from -.74 to -.80 (Miller et al., 2015).

***SSIS SEL Screening/Progress Monitoring Scales.*** Although the SSIS-PSG has demonstrated technical adequacy and usability among raters, it is not well aligned with the CASEL Five model or SEL intervention programs designed to teach students skills consistent with the five model components. The SSIS SEL Screening/Progress Monitoring Scales was developed to specifically measure all five CASEL components (i.e., self awareness, relationship

skills, responsible decision making, social awareness, and self management) for the purposes of (1) identifying students in need of social emotional intervention and (2) monitoring progress of student skill development during and after implementation of the intervention. These SEL skills were also contextualized with academic skills from the SSIS-PSG (i.e., Motivation to Learn, Reading, and Mathematics) to develop a comprehensive screening tool for use in schools (Elliot, Davies, Frey, & Cooper, in press). Additionally, there is preliminary research to support the technical adequacy and usability of the SSIS SEL Screening/Progress Monitoring Scales, which provides much promise for the integration of this screening instrument in schools.

The SSIS SEL Screening/Progress Monitoring Scales was developed over two stages. In the first stage of measurement development, three social behavior researchers worked to independently assign the social skills items from the SSIS Improvement System to a CASEL social emotional skill category. This activity allowed researchers to determine groups of objective behaviors that likely fit within each of the five social emotional domains. Elliott and colleagues (in press) reported an item assignment agreement of 61%, 65%, and 71% among the three potential groups of researchers. In the second round of review, researcher disagreement was discussed and a final consensus agreement for item assignment of 93% of items was achieved (i.e., 43 of 46 items). In the second stage of SSIS SEL Screening/Progress Monitoring Scales development, researchers conducted a Q-sort method with six teachers to provide further evidence for the alignment of the SSIS SEL Screening/Progress Monitoring Scales item content with the CASEL domains. Teacher naïve to the purpose of the SSIS SEL Screening/Progress Monitoring Scales and CASEL model were provided definitions for each of the social emotional domains and 25 randomly ordered slips of paper with each 5-level performance descriptor for each of the SSIS SEL Screening/Progress Monitoring Scales rubrics. Teachers worked



independently to assign each performance descriptor to a social emotional domain, and subsequently ordered the slips from lowest to highest level. Five of the six teachers completed the activity. Of the five teachers that completed the activity, all completed independently and were 100% correct (Elliott et al., in press).

The initial SSIS SEL Screening/Progress Monitoring Scales validation study provided promising preliminary evidence for the use of the measure in schools. As mentioned previously, results demonstrated that the content of the SSIS SEL Screening/Progress Monitoring Scales was well aligned with the CASEL Five model. Specifically, both teachers and experienced social skills highly agreed on assignment of skills and alignment the CASEL skill definitions, which suggests that the SSIS SEL Screening/Progress Monitoring Scales is valid sample or representation of CASEL skill domains. Additionally, consistent with extant research findings, the results demonstrated girls were consistently rated as higher on social and academic skills than boys. The SSIS SEL Screening/Progress Monitoring Scales was also sensitive to developmental level differences in students, with increases in scores as student progresses in age/grade. Further, the initial reliability estimates for the SSIS SEL Screening/Progress Monitoring Scales were high for the Social Emotional (SE) Composite of the five SEL scales (alpha .91), Academic Functioning (AF) Composite of the three academic criteria (alpha .90), and total SEL score (alpha .93) of all eight scales. Using these reliability estimates, Elliott and colleagues (in press) calculated the standard error of measure (SEM) for each subscale, which ranged from low (.34) to high (.41). Overall, reliability estimates obtained demonstrate preliminary highly reliable scores. The SSIS SEL Screening/Progress Monitoring Scales also accurately discriminated between at-risk and not at-risk students using PSG ratings. Specifically, the SSIS SEL Screening/Progress Monitoring Scales SE Composite accurately identified 60.5% of students as

true positives and 92.4% as true negatives. Additionally, the SSIS SEL Screening/Progress Monitoring Scales AF Composite yielded high sensitivity and specificity estimates, ranging from 86.7% and 84.4% of students as true positives and 91% and 83.9% of students as true negatives for SSIS-PSG Reading and Math, respectively. Lastly, participating teachers rated the SSIS SEL Screening/Progress Monitoring Scales as highly feasible, time efficient, and relevant. Although the SSIS SEL Screening/Progress Monitoring Scales initial validation study conducted by Elliott and colleagues (in press) indicated that the SSIS SEL Screening/Progress Monitoring Scales is well-aligned with CASEL framework, reliable, relevant, sensitive, and efficient, further research is required to replicate and extending findings to schools within the United States, as well as assess its application to more diverse sample of students and teachers. Additionally, further research should include repeated measures of the SSIS SEL Screening/Progress Monitoring Scales concurrent with other validated measures of social emotional functioning, such as the SSIS-PSG and SIBS/SEBS. This evidence will further understanding of SSIS SEL Screening/Progress Monitoring Scales use and application within US elementary, middle, and high schools.

### **Purpose and Research Questions**

Various methods exist for assessing students social and emotional skills, including direct observations, role-plays, and parent, teacher, and student interviews. More recently, however, clinicians and researchers have most frequently relied on rating scales for assessing social and emotional skills (Crowe, Beauchamp, Catroppa, & Anderson, 2011; Humphrey et al., 2011). Rating scales are relatively efficient tools for assessing individual social and emotional behaviors, and information provided by valid and reliable measures is particularly useful for determining the need for intervention services. Due to the intensity of academic, behavioral, and

social problems that students with emotional and behavioral problems exhibit, early identification and intervention services are integral to promoting student academic success and well-being. Extant research has clearly demonstrated the beneficial impact of early identification and intervention services, particularly with children and adolescents with emotional and behavioral problems. Additionally, establishing technical adequacy and utility of screening measures of social emotional functioning is vital in this process to ensure efficient and accurate identification and subsequent implementation of interventions targeting deficits of skills.

The central purpose of the present study is to further evaluate the technical adequacy and usability of the SSIS SEL Screening/Progress Monitoring Scales, which was designed to identify students in need of social emotional interventions to improve skills and monitor process of skill development during and following intervention. Specifically, the study will seek to: (a) assess the reliability of the SSIS SEL Screening/Progress Monitoring Scales; (b) evaluate the concurrent validity of the SSIS SEL Screening/Progress Monitoring Scales with well-researched screeners; (c) investigate the short-term predictive validity of the SSIS SEL Screening/Progress Monitoring Scales with other indicators behavioral and academic performance at school; (d) assess the classification accuracy of the SSIS SEL Screening/Progress Monitoring Scales; and (d) supplement extant social validity data on the SSIS SEL Screening/Progress Monitoring Scales, including the usability and feasibility as rated by teachers.

For these purposes, the present study investigated the following research questions:

1. Does the SSIS SEL Screening/Progress Monitoring Scales demonstrate adequate internal consistency and test-retest reliability?

2. Is the SSIS SEL Screening/Progress Monitoring Scales significantly correlated with well-established universal screening measures, including the SIBS/SEBS and SSIS-PSG?
3. Are the rates of detection of social emotional and behavioral risk by the SSIS SEL Screening/Progress Monitoring Scales similar to other behavior screening measures?
4. What is the overall risk classification accuracy (i.e., sensitivity, specificity, positive predictive power, and negative predictive power) of the SSIS SEL Screening/Progress Monitoring Scales as compared to the SIBS/SEBS and SSIS-PSG?
5. Are the results of the SSIS SEL Screening/Progress Monitoring Scales significantly correlated with indicators of school performance, including office discipline referrals (ODRs), suspension, attendance, and academic grades?
6. How do teachers rate the usability, acceptability, and feasibility of the SSIS SEL Screening/Progress Monitoring Scales?

## CHAPTER 2 METHOD

### Participants and Setting

Participants included teachers and students from general education classrooms in grades 1 through 4 in public elementary schools in southeastern Louisiana. Ten teachers from two public elementary schools consented to participate and completed screeners on all students within their classrooms. Participating teachers averaged 30.3 years of age ( $SD = 5.50$ , range = 25-44), and 6.8 years of experience in the education field ( $SD = 5.65$ , range = 3-21). All teachers were female ( $N = 10$ ). The majority of the sample identified as White, Non-Hispanic (60%,  $n = 6$ ), followed by African American (30%,  $n = 3$ ) and Hispanic/Latino (10%,  $n = 1$ ). Lower and upper elementary grade levels were largely equally represented from 1<sup>st</sup> to 4<sup>th</sup> grade. Self-reported teacher demographic data are displayed in Table 1.

Table 1  
Teacher Demographic Information

Category	<i>n</i>	%	Category	<i>n</i>	%
Sex			Race/Ethnicity		
Female	10	100%	White, Non-Hispanic	6	60%
Male	0	0%	African American	3	30%
			Hispanic/Latino	1	10%
Age			Asian American	0	0%
21-25	1	10%	Native American	0	0%
26-30	5	50%	Other	0	0%
31-35	2	20%	Grade Taught		
36-40	0	0%	1 <sup>st</sup>	3	30%
40+	1	10%	2 <sup>nd</sup>	2	20%
Years Teaching			3 <sup>rd</sup>	2	20%
0-4 years	6	60%	4 <sup>th</sup>	3	30%
5-9 years	2	20%			
10-14 years	1	10%			
15+ years	1	10%			

A total of 192 students comprised the sample at the outset of the study. One of the 10 teachers did not complete the student-level data sheet, and an additional 6 students who were initially enrolled in the study (3.1% of the original sample) withdrew from the participating schools during the semester. Therefore, 186 students completed the study by remaining enrolled in the participating schools for the duration of the semester, and student-level data was obtained from 165 students at the end of the fall semester.

Student demographic information was collected from each participating teacher. Eighty-eight students were female (45.8%), and 104 were male. Age of participating students ranged from 6 to 10, with an average age of 7.84 years ( $SD = 1.35$ ). However, age was only reported for approximately half of the overall sample (49%,  $n = 94$ ). Consistent with teacher participation, younger elementary grade levels (48.4% in 1<sup>st</sup> and 2<sup>nd</sup> grades,  $n = 93$ ) and upper elementary grade levels (51.6% in 3<sup>rd</sup> and 4<sup>th</sup> grades,  $n = 99$ ) were largely consistently represented. Majority of students were identified as African American (65.1%,  $n = 125$ ), with the remainder of the sample being comprised of Hispanic/Latino (24.0%,  $n = 46$ ), White, Non-Hispanic (4.2%,  $n = 8$ ), Asian American (2.6%,  $n = 5$ ), and Other (3.1%,  $n = 6$ ).

Additional student data was obtained regarding educational programming and accommodations. Fourteen students (7.3% of the sample) had a 504 plan for educational accommodations, and 16 students (8.3% of the sample) were classified as special education under one or more of the 13 special education disability categories in IDEIA. Teacher-reported student demographic data are located in Table 2.

Table 2  
Student Demographic Information

Category	<i>n</i>	%	Category	<i>n</i>	%
----------	----------	---	----------	----------	---

Sex  
(table cont'd.)

Race/Ethnicity

Category	<i>n</i>	%	Category	<i>n</i>	%
Male	104	54.2%	White, Non-Hispanic	8	4.2%
Female	88	45.8%	African American	125	65.1%
Grade			Hispanic/Latino	46	24.0%
			Asian American	5	2.6%
			Native American	0	0.0%
			Other	6	3.1%
Age			Special Education		
			General Education	176	91.7%
			Special Education	16	8.3%
6	21	10.9%	504 Plan		
7	23	12.0%			
8	8	4.2%			
9	34	17.7%			
10	8	4.2%	No 504 Plan	178	92.7%
			504 Plan	14	7.3%

## Measures

**Teacher Demographic Information Form.** Participating teachers completed the Teacher Demographic Information Form, which was created for this study. Items on the form assessed age, gender, race/ethnicity, current grade taught, and years of teaching experience. A copy of the Teacher Demographic Information Form can be found in Appendix A, and the data obtained using this form is presented in Table 1 above.

**Student Demographic Information Form.** Participating teachers also completed a brief Student Demographic Information Form on each student in their class. Items on the form assessed age, gender, race/ethnicity, current grade level, special education classification, and 504 accommodations. A copy of the Student Demographic Information Form can be found in Appendix B, and the data obtained using this form is presented in Table 2 above.

**SSIS SEL Screening/Progress Monitoring Scales.** The SSIS SEL Screening/Progress Monitoring Scales (Elliott & Gresham, 2017) is a screening measure designed to efficiently assess student social emotional learning and academic skills in eight domains. Teachers evaluate

students relative to expectations in the following five social emotional skills and three academic behavior domains: Self-Awareness, Relationship Skills, Responsible Decision-Making, Social Awareness, Self-Management, Motivation to Learn, Reading Skills, and Mathematic Skills. The SSIS SEL Screening/Progress Monitoring Scales provides scores for each of the eight domains, with scores ranging from 1 to 5. The SSIS SEL Screening/Progress Monitoring Scales also provides a Social Emotional (SE) Composite score ranging from 5 to 25 and an Academic Functioning (AF) Composite score ranging from 3 to 15. Elliott et al. (in press) purported that for screening and intervention planning purposes, scores of 1 and 2 on a particular domain are indicative of at-risk. Additionally, students with an SE Composite score of 10 or less are considered at-risk socially, and students with an AF composite of 6 or less are considered at-risk academically.

Elliott et al. (in press) provided preliminary validity and reliability evidence for the SSIS SEL Screening/Progress Monitoring Scales with elementary students in Australia. Results suggest acceptable internal consistency estimates for combined eight domains ( $\alpha = .93$ ), SE Composite of the five social emotional domains ( $\alpha = .91$ ) and AF Composite of the three academic domains ( $\alpha = .93$ ), which resulted in low standard errors of measurement for each of the eight subscales. Using the SSIS-PSG Prosocial Scale as comparison, the SSIS SEL Screening/Progress Monitoring Scales SE Composite accurately identified 60.5% of students as true positives and 92.4% as true negatives. Additionally, the SSIS SEL Screening/Progress Monitoring Scales AF Composite yielded high sensitivity and specificity estimates, ranging from 86.7% and 84.4% of students as true positives and 91% and 83.9% of students as true negatives for SSIS-PSG Reading and Math, respectively.



In the current study, participating teachers completed the SSIS SEL Screening/Progress Monitoring Scales on all students in their classrooms to assess student social emotional and academic skills. The SSIS SEL Screening/Progress Monitoring Scales was administered twice during the study, once at approximately 7 weeks into the academic year concurrently with completion of the other screeners (i.e., combined SIBS/SEBS, SSIS-PSG, and SSIS SEL Core Skills Scale) and again 6 weeks following the first administration.

***Student Internalizing Behavior Screener.*** The Student Internalizing Behavior Screener (SIBS) is a teacher-completed universal screening instrument designed to measure internalizing behavior problems in students (Cook et al., 2011). The SIBS is comprised of the following seven behavioral indicators: nervous/worried or fearful, bullied by peers, spends time alone, clings to adults, withdrawn, seems sad or unhappy, and complains about being sick or hurt. Teachers rate students utilizing a 4-point Likert response scale to indicate their perception of frequency of observed target behaviors (i.e., 0 = *Never*, 1 = *Rarely/Seldom*, 2 = *Occasionally/Moderately*, 3 = *Frequently/Almost Always*). Scores are derived from calculating a total sum of all SIBS items. Higher total scores on the SIBS suggest a greater frequency of teacher-perceived internalizing behavior problems (i.e., total score range of 0 to 21). The following categories and cutoff scores are provided for interpretation: at-risk (8 or more), on the radar (4 to 7), and not at risk (0 to 3) (Cook et al., 2011).

Previous studies have supported the reliability and validity of the SIBS. Cook and colleagues (2011) demonstrated that the SIBS maintains adequate internal consistency ( $\alpha = .81$  and  $.79$ ) and test-retest reliability ( $r = .74$ ). Acceptable convergent validity of the SIBS was also demonstrated using the ASEBA *Teacher Report Form* Internalizing scale ( $r = .82$ ).

***Student Externalizing Behavior Screener.*** The Student Externalizing Behavior Screener (SEBS) is a universal screening instrument designed to measure externalizing behavior problems (Cook et al., 2012). The SEBS is comprised of the following 7 items: defiant or oppositions to adults, fights or argues with peers, bullies others, gets angry easily, lies to get out of trouble, disrupts class activities, and has difficulty sitting still. As on the SIBS, each item is arranged on a 4-point Likert response scale ranging from *Never* to *Frequently/Almost Always*. Higher total scores on the SEBS are indicative of a greater frequency of externalizing behavior problems in students. SEBS scores range from 0 to 21 and can be interpreted across three risk categories: at-risk (9 or higher), on the radar (5 to 8), and not at-risk (0 to 4) (Cook et al., 2012).

Research on the SEBS revealed excellent internal consistency estimates ( $\alpha = .89$  and  $.84$ ) and strong test-retest reliability ( $r = .92$ ; Cook et al., 2012). Furthermore, a strong correlation of SEBS scores with the ASEBA *Teacher Report Form* Externalizing scale ( $r = .87$ ) and with the SRSS ( $r = .91$ ; Drummond, 1994) provided convergent validity evidence.

As a well-researched and technically sound universal screening measure, the SIBS and SEBS served as the criterion measure in the current study. Teachers completed the SEBS and SIBS concurrently on all students in their classroom to assess emotional and behavioral risk related to internalizing and externalizing behaviors at approximately 7 weeks into the academic school year. In the combined format, items from the SEBS were presented first (#1-7), followed by items from the SIBS (#8-14). The combined SEBS and SIBS measure is located in Appendix C in the same format completed by teachers in the study.

***SSIS Performance Screening Guide.*** The Social Skills Improvement System – Performance Screening Guide (SSIS-PSG; Elliott & Gresham, 2007) is a universal screening measure designed to assess student academic and behavioral skills in four performance domains:

Prosocial Behavior, Motivation to Learn, Reading Skills, and Math Skills. The SSIS-PSG provides behaviorally anchored, multi-level descriptors for each of the performance areas. Teachers assign a global rating from 1 (substantial risk or high concern), 2 or 3 (moderate concern or difficulty), and 4 or 5 (average to above average) for each performance domains based on their observations and interactions with students in their classroom. For the current study, a student is considered at-risk or in need of additional intervention supports if he or she is assigned a rating of 2 or lower.

In extant research conducted by Elliott and Gresham (2007) on the SSIS-PSG, test-retest reliability coefficients ranged from  $r = .68$  to  $r = .74$  and interrater reliability coefficients were between  $r = .55$  and  $r = .68$  across skill areas for a sample of elementary-aged students, suggesting strong to moderate evidence of reliability of the SSIS-PSG.

In the current study, the SSIS-PSG was completed by all participating teachers approximately 7 weeks into the academic year. In addition to the combined SIBS/SEBS, all analyses were also conducted utilizing the SSIS-PSG as criterion to replicate the Elliott et al. (in press) study.

***SSIS SEL Edition-Rating Forms (SSIS SEL RF).*** The SSIS SEL Edition—RF is a reconfiguration of the SSIS-RS to fit within the five CASEL domains (Gresham & Elliott, 2017). The SSIS SEL RF is comprised of the following five domains: Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision Making. Additionally, the SSIS SEL RF *Teacher Form* is comprised of the five SEL domains and an Academic Competence domain, consisting of 7 items corresponding to academic functioning. SSIS SEL RF items are rated on a 4-point Likert scale to indicate the frequency of observed behavior, ranging from *Never* to *Always*. Scores on the SSIS SEL RF subscales are expressed as

standard scores ( $M = 100$   $SD = 15$ ). The SSIS SEL RF provides an SEL Composite, 5 SEL competencies, Core Skills, and Academic Competence scores. The Core Skills Scale contains 10 items that directly correspond to skills taught in the Core Skills units of the SEL Classwide Intervention Program.

For the present study, teachers completed the Core Skills scale for each participating student. The SSIS SEL Core Skills measure is presented in Appendix D in the same format completed by teachers in the study.

**Usage Rating Profile – Assessment.** The *Usage Rating Profile – Assessment* (URP-A; Chafouleas, Miller, Briesch, Neugebauer, & Riley-Tillman, 2012) is a 28-item self-report measure designed to evaluate perceived acceptability, feasibility, and usability of an assessment instrument. The URP-A items are rated on a 6-point Likert scale indicating the degree of agreement, from *Strongly Disagree* to *Strongly Agree*. The URP-A assesses six domains of assessment usability, including Acceptability (appropriateness of the assessment and interest in its use), Understanding (knowledge of assessment and its procedures), Home-School Collaboration (perceived necessity of collaboration in completing the assessment), Feasibility (ease of use), System Climate (extent of fit within school systems), and System Support (need for additional support to use the assessment). Internal consistency estimates of the six URP-A domains range from acceptable to high ( $\alpha = .63$  to  $\alpha = .90$ ). Total scores on the URP-A reflect overall perceptions of the usability of an assessment. The Total Usability score and the six domain scores are interpreted using mean item scores.

The URP-A was adapted for use in the current study to include only the 15 items comprising the Acceptability and Feasibility subscales. The URP-A was administered to teachers following the second administration of the SSIS SEL Screening/Progress Monitoring Scales. The

adapted URP-A is presented in Appendix E in the same format completed by teachers in the study.

***Student-level Data.*** Student-level data was collected as an additional measure of short-term predictive validity of the SSIS SEL Screening/Progress Monitoring Scales with other academic and behavioral performance indicators at school. A student-level data sheet was developed for the current study, and provided to each teacher with corresponding student codes. Participating teachers reported student letter grades in core academic subjects only (i.e., Math and English/Language Arts). Letter grades were then assigned a point value (i.e., A = 4.0, B = 3.0, C = 2.0, D = 1.0, and F = 0.0). Additionally, teachers provided an estimation of ODRs, suspensions, and absences for all students in their respective classrooms from the beginning of the current school semester to the time of third administration. The student-level data sheet is presented in Appendix F in the same format completed by teachers in the study.

## **Procedure**

***Administrator and Teacher Consent.*** Administrative informed consent was obtained from two public elementary schools. Teachers were then recruited via in-person contact and informed of the study procedures. Informed consent was obtained from 10 teachers for completing screening measures on all students in their respective classrooms. Contact information was also obtained from each teacher. Teachers were offered a nominal incentive for their participation. Upon completing all data collection phases, participating teachers received a \$10 gift card.

***Teacher Training.*** After teachers were recruited, a brief training was conducted to provide an overview and directions regarding completion of the coding spreadsheet, the various screening measures, and study procedures. The researcher reviewed the coding process, and

provided an example of the coding spreadsheet to be completed at Time 1. The researcher then demonstrated how to password protect the coding spreadsheet and provided an email address to forward each coding document to at the conclusion of the training session. Teachers independently created a personal password for their corresponding coding spreadsheet, in which the researcher did not access. Additionally, the researcher reviewed the instructions for each of the screening measure, and provided teachers with the study materials for Time 1 administration at the completion of the training session (i.e., Teacher Demographic Information form, Student Demographic Information form, SSIS SEL Screening/Progress Monitoring Scales, the combined SIBS/SEBS, SSIS-PSG, SSIS SEL Core Skills Scale). Any questions or concerns regarding the study procedures or screening instruments were addressed.

**Data Collection.** Data collection occurred in three phases during the fall semester (Time 1, Time 2, Time 3 administration). Participating teachers were provided with binders containing all screening instruments for Time 1 at the teacher training session. Before completing the screening measures, teachers were provided with codes and instructed to fill in names of each student within their classroom. The codes and corresponding identifying information were maintained in a password-protected spreadsheet. Each teacher created an individual password and locked the document prior to forwarding to the researcher to be maintained until Time 2 administration. Teachers were then instructed to complete the Teacher Demographic Information Form, Student Demographic Information Form, SSIS SEL Screening/Progress Monitoring Scales, SSIS-PSG, SIBS/SEBS, and the SSIS-SEL Core Skills on all students in their classroom at approximately 7 weeks into the school year. This elapse in time allowed teachers to become familiar with students in their respective classrooms and likely increased the likelihood of accurate behavior ratings. In completing the screeners, teachers used individualized student

codes rather than student names to ensure that confidentiality of student identity was maintained. Researchers collected the completed screeners within two weeks of Time 1 administration.

Approximately 6 weeks after the initial screening, the SSIS SEL Screening/Progress Monitoring Scales and URP-A was provided to each teacher to be completed for Time 2 administration. The password-protected teacher coding spreadsheet was forwarded to each teacher to ensure accuracy of reporting during Time 2 administration. Teachers completed the URP-A regarding their perceptions of the usability of the SSIS SEL Screening/Progress Monitoring Scales at Time 2. Teachers were allotted one week to complete the SSIS SEL Screening/Progress Monitoring Scales and URP-A at Time 2 administration. At the end of the fall school semester, the Student-Level Data Sheet was provided to each participating teacher to be completed for Time 3 administration. As with Time 2, the password-protected spreadsheet was forwarded to all participating teachers to assist in completing the form. Researchers collected the completed Student-Level Data Sheet approximately one week following Time 3 administration.

## CHAPTER 3 RESULTS

### Descriptive Statistics

Descriptive statistics are provided in Table 3 to summarize data on screening measures completed on students in the current sample ( $N = 192$ ). Results of the SSIS SEL Screening/Progress Monitoring Scales demonstrated a mean Social Emotional (SE) Composite score of 17.50 with a standard deviation of 4.75, and a mean Academic Functioning (AF) Composite score of 9.56 with a standard deviation of 3.35. The SSIS SEL Screening/Progress Monitoring Scales subscale means ranged from 3.43 ( $SD = 1.13$ ) to 3.63 ( $SD = 1.05$ ), as presented in Table 3 below. Results of the SSIS-PSG revealed a Prosocial Behavior mean rating of 3.77 ( $SD = 1.05$ ), and a Motivation to Learn domain mean of 3.67 ( $SD = 1.19$ ). The combined SIBS/SEBS resulted in a total mean score of 7.98 with a standard deviation of 8.56. Lastly, the mean rating from teachers on the SSIS Core Skills Scale was 30.22 with a standard deviation of 6.69.

Table 3  
Descriptive Statistics for Screening Measures

	<i>M</i>	<i>SD</i>	Range of Scores
SSIS SEL SE Composite	17.50	4.75	5-25
Self-Awareness	3.46	1.02	1-5
Relationship Skills	3.52	1.07	1-5
Responsible Decision-Making	3.63	1.05	1-5
Social Awareness	3.46	1.07	1-5
Self-Management	3.43	1.13	1-5
SSIS SEL AF Composite	9.56	3.35	3-15
Motivation to Learn	3.61	1.21	1-5
Reading Skills	3.13	1.30	1-5
Mathematic Skills	3.17	1.30	1-5
SSIS-PSG Prosocial Behavior	3.77	1.05	1-5



SSIS-PSG Motivation to Learn	3.67	1.19	1-5
SIBS/SEBS	7.98	8.56	0-32
SIBS Alone	2.55	3.60	0-17
SEBS Alone	5.43	6.36	0-21
SSIS Core Skills	30.22	6.69	10-40

Student-level data are presented in Table 4. At the end of the school semester, teachers completed the Student-Level Data form on all students in their classroom, indicating the estimated number of office discipline referrals (ODRs), suspensions, absences, and letter grades in ELA and math. Student-level data was collected on 165 students from nine classrooms. ODRs ranged from 0 to 24, with a mean of 0.83 and standard deviation of 3.10. Of the 165 students in the final sample, 27 total students accounted for all reported ODRs, approximately 16% of the overall sample of students at Time 3. Of the 27 students, 18 reportedly received between 1 to 4 ODRs. Three students earned an estimated 5 to 9 ODRs, and 6 students reportedly received 10 or more ODRs during the school semester. Reported suspensions ranged from 0 to 3 ( $M = 0.06$ ,  $SD = 0.33$ ), with 7 students accounting for all teacher-reported suspensions. Per teacher report, the mean of days absent was 3.73, with a standard deviation of 5.18. Lastly, estimated letter grades for ELA and Math were collected and transformed for correspond to a 4.0 GPA scale. The overall mean GPA for the current study sample was 2.40 ( $SD = 1.32$ ), which is equivalent to a C letter grade.

Table 4  
Descriptive Statistics for Student-Level Data

	<i>M</i>	<i>SD</i>	Range
Office Discipline Referrals (ODRs)	0.83	3.10	0-24
Suspensions	0.06	0.33	0-3
Absences	3.73	5.18	0-35
Academic Grades (GPA)	2.40	1.32	0.0-4.0

ELA GPA	2.47	1.31	0.0-4.0
Math GPA	2.31	1.32	0.0-4.0

### Reliability of the SSIS SEL Screening/Progress Monitoring Scales

Analyses were conducted on scores from both Time 1 and Time 2 administrations to assess internal consistency and test-retest reliability of the SSIS SEL Screening/Progress Monitoring Scales. The consistency of item responses on the SSIS SEL Screening/Progress Monitoring Scales was assessed by evaluating the correlations among items using Cronbach's coefficient alpha as the estimate of internal consistency. Internal consistency was calculated for the SSIS SEL Screening, as well as the SE and AF Composite scales at Time 1 and Time 2 administration. As noted previously, extant research indicates that levels of internal consistency may be considered adequate when reaching .7 or .8 (Field, 2009). As presented in Table 5, teachers' ratings on the SSIS SEL Screening/Progressing Monitoring Scales yielded Cronbach's  $\alpha$  of .93 (Time 1) and .96 (Time 2), demonstrating adequate internal consistency reliability for the overall scale at both Time 1 and Time 2 administrations. Additionally, internal consistency estimates for the SE Composite scale were .93 and .96 at Time 1 and Time 2, respectively. The AF Composite scale resulted in Cronbach's  $\alpha$  of .87 (Time 1) and 0.91 (Time 2). These estimates indicate that the SE and AF Composite scales also demonstrate adequate internal consistency reliability in comparison to recommended reliability coefficients (Field, 2009).

Table 5  
Internal Consistency Reliability of the SSIS SEL Screening/Progress Monitoring Scales

	Time 1	Time 2
SSIS SEL Screening	$\alpha = .93$	$\alpha = .96$
SE Composite Scale	$\alpha = .93$	$\alpha = .96$
AF Composite Scale	$\alpha = .87$	$\alpha = .91$

The stability of scores or test-retest reliability was calculated using Pearson's  $r$  coefficient for the SSIS SEL Screening/Progress Monitoring Scales scores at Time 1 and Time 2 administration. Extant literature suggests that reliability coefficients be at least .70 for measures used for screening purposes (Reynolds & Livingston, 2014). The ratings on the SSIS SEL Screening/Progress Monitoring Scales SE Composite scale resulted in a reliability coefficient of .68, with reliability of subscales ranging from .61 to .68. Teacher ratings on the AF Composite scale produced a reliability coefficient of .58, with reliability of subscales ranging from .56 to .64. Test-retest reliability results are located in Table 6.

Table 6  
Test-Retest Reliability of the SSIS SEL Screening/Progress Monitoring Scales

	$r$
SSIS SEL SE Composite	.68*
Self-Awareness	.61*
Relationship Skills	.61*
Responsible Decision-Making	.65*
Social Awareness	.61*
Self-Management	.62*
SSIS SEL AF Composite	.58*
Motivation to Learn	.60*
Reading Skills	.64*
Math Skills	.56*
*Pearson correlation is significant, $p < .01$ .	

### Concurrent Validity

To determine concurrent validity of the SSIS SEL Screening/Progress Monitoring Scales, the relationship among screener scores and identification rates were evaluated using Time 1 data. The SSIS-PSG, SIBS, and SEBS scores were used as the primary criterion measures given the extensive research supporting the technical adequacy of these measures. The association between

scores was assessed using Pearson's  $r$  correlation coefficients. The strength of the resulting bivariate correlations was interpreted using the following recommendation: less than .30 are small, .30 to .49 are moderate, and greater than .50 are strong (Cohen, 1977).

As indicated in Table 7, the SSIS SEL Screening/Progress Monitoring Scales SE Composite and combined SIBS/SEBS scores resulted in a significant negative correlation coefficient of  $-.68$ , indicating that as students scored higher on social emotional skills, they tended to score lower on internalizing and externalizing behavior problems. The SSIS SEL Screening/Progress Monitoring Scales SE Composite score was significantly related to behavioral domains of the SSIS-PSG,  $r = .58$  on Prosocial Behavior and  $r = .66$  on Motivation to Learn. The greater number and/or frequency of reported social skills on the SSIS SEL Screening/Progress Monitoring Scales, the higher the rating on Prosocial Behavior and Motivation to Learn. Additionally, the SSIS SEL Screening/Progress Monitoring Scales AF Composite resulted in a small negative correlation coefficient for the combined SIBS/SEBS ( $r = -.24, p < .01$ ) and strong positive correlation coefficients for the SSIS-PSG Prosocial Behavior ( $r = .67, p < .01$ ) and Motivation to Learn ( $r = .72, p < .01$ ) domains. Overall, the SSIS SEL Screening/Progress Monitoring Scales SE and AF Composite scales were found to be strongly and significantly correlated with other well-established universal behavior and social emotional screeners, providing support for concurrent validity.

Other associations among screening measures were also found to be significantly strong. Significant positive correlations were demonstrated between the SSIS SEL Core Skills Scale and the SSIS-PSG Prosocial Behavior ( $r = .53, p < .01$ ) and Motivation to Learn ( $r = .48, p < .01$ ) domains. The SSIS SEL Core Skills Scale combined SIBS/SEBS resulted in a significant negative correlation coefficient of  $-.63$ . Additionally, the combined SIBS/SEBS resulted in

moderate negative correlations with the SSIS-PSG Prosocial Behavior ( $r = -.44, p < .01$ ) and Motivation to Learn ( $r = .50, p < .01$ ). These results are also presented in Table 7.

Table 7  
Correlations Between Universal Behavioral and Social Emotional Screening Scores

	SSIS SEL SE Composite	SSIS SEL AF Composite	SSIS-PSG Motivation to Learn	SSIS-PSG Prosocial Behavior	SIBS/SEBS	SSIS SEL Core Skills Scale
SSIS SEL SE Composite		.63*	.66*	.58*	-.68*	.65*
SSIS SEL AF Composite			.72*	.67*	-.31*	.40*
SSIS-PSG Motivation to Learn				.63*	-.50*	.48*
SSIS-PSG Prosocial Behavior					-.44*	.53*
SIBS/SEBS						-.63*
SSIS SEL Core Skills Scale						

\*Pearson correlation is significant at .01.

### Conditional Probabilities and ROC Analysis

Conditional probabilities were computed to assess the accuracy of risk classification of the SSIS SEL Screening SE Composite scores using the combined SIBS/SEBS as the criterion measure and the proposed risk classification cutoff score of 10 (Elliott et al., in press). Extant research suggests that a cutoff of .60 or greater should be used when interpreting the classification accuracy indices (Shapiro, Keller, Edwards, Lutz, & Hintze, 2006). Table 8

provides a summary of the classification accuracy of the SSIS SEL Screening/Progress Monitoring Scales.

In evaluating the classification accuracy of the SSIS SEL Screening/Progress Monitoring Scales SE Composite according to the combined SIBS/SEBS, the sensitivity index was .286, indicating the probability that a student was correctly identified by the SSIS SEL Screening as at-risk compared to all students identified as at-risk by the combined SIBS/SEBS. Otherwise stated, only 28.6% of students identified by the criterion measure as at-risk ( $n = 63$ ) were also identified by the SSIS SEL Screening ( $n = 21$ ). The specificity index was 0.977, demonstrating the proportion of students who were correctly identified by the SSIS SEL Screening SE measure as not at-risk ( $n = 171$ ) compared to all the students who were identified as not at-risk according to the SIBS/SEBS ( $n = 129$ ). The positive predictive power was .857, indicating that 85.7% of students detected as at-risk on the SSIS SEL Screening were also identified by the SIBS/SEBS. The negative predictive power was .737, which demonstrates that 73.7% of students identified as not at-risk on the SSIS SEL Screening were also not identified as at-risk by the SIBS/SEBS. Utilizing the previously established cutoff score and the SIBS/SEBS as the criterion measure, the SSIS SEL Screening SE Composite performed well in terms of specificity, positive predictive power, and negative predictive power. However, the sensitivity index was significantly lower than recommended cutoff of .60.

In addition, to replicate the Elliott et al. (in press) study, conditional probability indices were calculated for the SSIS SEL Screening SE Composite using the SSIS-PSG Prosocial domain as criterion. As with the SIBS/SEBS, the SSIS SEL Screening SE Composite scores resulted in inadequate sensitivity (.417) and positive predictive power (.476) using the SSIS-PSG Prosocial domain as criterion. Consistent with the combined SIBS/SEBS criterion, the specificity

(.935) and negative predictive power (.91) exceeded the recommended cutoff of 0.60. Results are presented in Table 8.

Table 8  
Summary of Classification Accuracy of SSIS SEL Screening SE Composite by the SIBS/SEBS and SSIS-PSG Prosocial

	SIBS/SEBS	SSIS-PSG Prosocial
Sensitivity	28.6%	41.7%
Specificity	97.7%	93.5%
Positive Predictive Power	85.7%	47.6%
Negative Predictive Power	73.7%	91.8%

Additionally, the screening measures were examined regarding consistency in rates of risk identification. Risk was evaluated utilizing previously defined cutoff scores for classification (i.e., SIBS/SEBS: Cook et al., 2011, 2012; Hartman et al., 2017; SSIS SEL Screener: Elliott et al., in press; and SSIS-PSG: Elliott & Gresham, 2007). Results are depicted in Table 9. The SSIS SEL Screening/Progress Monitoring Scales SE Composite identified 10.9% of the current sample ( $n = 21$ ) as at-risk for social and emotional concerns, and the AF Composite identified 22.9% of students ( $n = 44$ ) as at-risk academically. The combined SIBS/SEBS detected 63 students as at-risk, which corresponds to 32.1% of the sample. Additionally, the SSIS-PSG identified 32 students as at-risk in Motivation to Learn (16.7%) and 24 students as at-risk in Prosocial Behavior (12.5%). A Cochran's  $Q$  test was run to determine whether there were statistically different identification rates across screening measures. Results demonstrate that the SSIS SEL Screening AF and SE Composites, combined SIBS/SEBS, and SSIS-PSG behavioral domains identified significantly different numbers of students at-risk,  $Q(4) = 59.939, p < .01$ .

Table 9  
Rates of Risk Identification for Screening Measures

	SSIS SEL Screening SE Composite	SSIS SEL Screening AF Composite	SIBS/SEBS	SSIS-PSG Motivation to Learn	SSIS-PSG Prosocial Behavior
Number Identified At-Risk (N = 192)	21	44	63	32	24
Percentage Identified At-Risk	10.9%	22.9%	32.1%	16.7%	12.5%

In addition, rates of risk identification were also calculated for the SSIS SEL Screening/Progress Monitoring Scales SE scales. As presented in Table 10, the percentage of students identified as at-risk according to the SE scales cutoff score ranged from 14.6% (Responsible Decision-Making) to 22.4% (Self-Management) of the overall sample. Overall, the number and percentage of students identified were largely consistent across the five social emotional domains. A Cochran's  $Q$  test was run to determine whether there were statistically different identification rates across SSIS SEL Screening scales. Results demonstrate that the Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision-Making scales identified significantly different numbers of students at-risk,  $Q(4) = 11.193, p < .05$ .

Table 10  
Rates of Risk Identification for the SSIS SEL Screening SE Scales

	Self- Awareness	Self- Management	Social Awareness	Relationship Skills	Responsible Decision- Making
Number Identified At- Risk (N = 192)	37	43	38	32	28
Percentage Identified At- Risk	19.3%	22.4%	19.8%	16.7%	14.6%



Given the low sensitivity of the SSIS SEL Screening/Progress Monitoring Scales SE Composite as compared to the SIBS/SEBS and SSIS-PSG, a ROC analysis using the SEBS/SEBS as a criterion variable was calculated to determine the best cut score for the current sample to maximize sensitivity and specificity. Figure 1 illustrates the ROC Area Under the Curve (AUC) test for the SSIS SEL Screening SE Composite scale when predicting membership in the at-risk group. The AUC is a measure of how well a measure separates the sample into two groups, those with and without social emotional difficulties. An AUC result of .9 to 1 is considered an excellent test, while an area of .50 is considered to be at chance discrimination. The area under the SSIS SEL Screening SE Composite ROC curve is .84,  $p < .01$ , which is considered to be a good discrimination percentage. In addition, as presented in Figure 2, the ROC curve analysis revealed that a score of 15 was deemed to be the optimal cutoff point to maximize sensitivity and specificity.

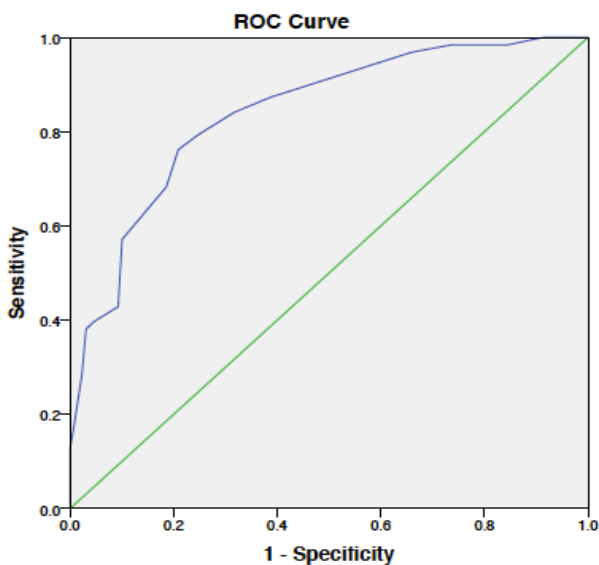


Figure 1. ROC Area Under the Curve (AUC) for the SSIS SEL Screening SE and combined SIBS/SEBS

**Coordinates of the Curve**

Test Result Variable(s): SEComposite1

Positive if Less Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
4.000	.000	.000
6.000	.036	.000
7.500	.091	.000
8.500	.145	.000
9.500	.200	.007
10.500	.327	.022
11.500	.418	.036
12.500	.436	.051
13.500	.473	.095
14.500	.618	.109
15.500	.745	.190
16.500	.800	.226
17.500	.818	.270
18.500	.836	.350
19.500	.855	.423
20.500	.964	.679
21.500	.982	.752
22.500	.982	.810
23.500	.982	.854
24.500	1.000	.920
26.000	1.000	1.000

The test result variable(s): SEComposite1 has at least one tie between the positive actual state group and the negative actual state group.

Figure 2. ROC Coordinates of the Curve for the SSIS SEL Screening SE using the combined SIBS/SEBS as Criterion

To replicate the Elliott et al. (in press) study, a ROC analysis on the SSIS SEL Screening SE Composite scale using the SSIS-PSG Prosocial domain as criterion was conducted. Figure 3 illustrates the ROC Area Under the Curve (AUC) test for the SSIS SEL Screening SE Composite scale when predicting membership in the at-risk group. The area under the SSIS SEL Screening SE Composite ROC curve using the SSIS-PSG Prosocial domain as criterion is .87,  $p < .01$ , which is considered to be a good discrimination percentage. In addition, as presented in Figure 4, the ROC curve analysis revealed that a score of approximately 15 was deemed to be the optimal cutoff point to maximize sensitivity and specificity. This is consistent with the cutoff score revealed by the ROC analysis using the combined SIBS/SEBS as criterion.

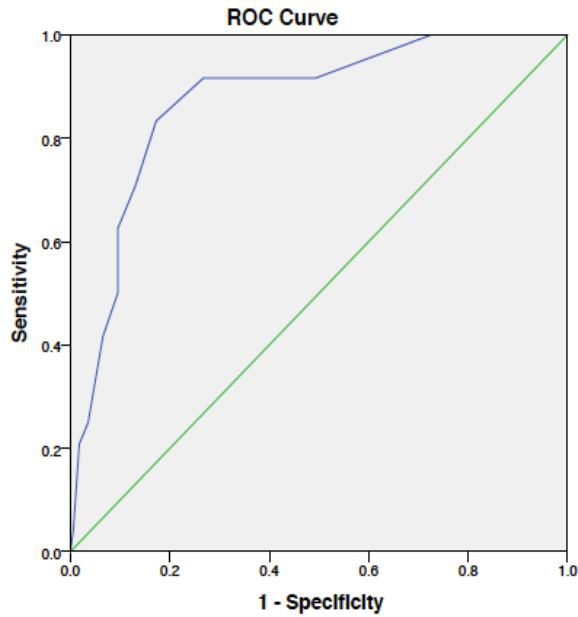


Figure 3. ROC AUC for the SSIS SEL Screening SE and SSIS-PSG Prosocial Domain

**Coordinates of the Curve**

Test Result Variable(s): SEComposite1

Positive if Less Than or Equal To <sup>a</sup>	Sensitivity	1 - Specificity
4.000	.000	.000
6.000	.042	.006
7.500	.125	.012
8.500	.208	.018
9.500	.250	.036
10.500	.417	.065
11.500	.500	.095
12.500	.625	.095
13.500	.708	.131
14.500	.833	.173
15.500	.917	.268
16.500	.917	.315
17.500	.917	.357
18.500	.917	.429
19.500	.917	.494
20.500	1.000	.726
21.500	1.000	.792
22.500	1.000	.839
23.500	1.000	.875
24.500	1.000	.935
26.000	1.000	1.000

The test result variable(s): SEComposite1 has at least one tie between the positive actual state group and the negative actual state group.

Figure 4. ROC Coordinates of the Curve for the SSIS SEL Screening SE and the SSIS-PSG Prosocial Domain

Following the ROC analysis on the SSIS SEL Screening SE Composite scale, conditional probabilities were again computed using the new cutoff score of 15 and the SIBS/SEBS as the criterion variable. Table 11 provides an updated summary of the classification accuracy of the SSIS SEL Screening/Progress Monitoring Scales by the combined SIBS/SEBS criterion measure. The sensitivity index increased from .175 to .683, which is within the range of the recommended cutoff for conditional probability indices. The specificity index was .814, which indicates that the SSIS SEL Screening SE Composite adequately identifies those not at-risk for social emotional deficits according to the combined SIBS/SEBS criterion. The positive predictive power was .642, and the negative predictive power was .84. Overall, the SSIS SEL Screening SE Composite performed well in terms of classification accuracy with the combined SIBS/SEBS as the criterion and the updated cutoff score of 15, with all indices meeting the recommended cutoff of .60.

Table 11  
Updated Summary of Classification Accuracy of SSIS SEL Screening SE Composite with cutoff of 15 by the SIBS/SEBS and SSIS-PSG Prosocial

	SIBS/SEBS	SSIS-PSG Prosocial
Sensitivity	68.3%	91.7%
Specificity	81.4%	73.2%
Positive Predictive Power	64.2%	32.8%
Negative Predictive Power	84.0%	98.4%

The risk identification rate for the SSIS SEL Screening SE Composite scale was also computed utilizing the new cutoff score of 15. Table 12 provides an updated summary of the identification rates of the SSIS SEL SE Screening and other universal screener scores. Once the SSIS SEL Screening SE Composite cutoff score was adjusted, the number and percentage of

students identified as at-risk increased, and was more consistent with the SIBS/SEBS. A Cochran's  $Q$  test was run to determine whether there were statistically different identification rates across screening measures following the updated cut score. Results demonstrate that the SSIS SEL Screening AF and SE Composites, combined SIBS/SEBS, and SSIS-PSG behavioral domains identified significantly different numbers of students at-risk,  $Q(4) = 64.862, p < .01$ .

Table 12  
Updated Summary of Rates of Risk Identification for Screening Measures

	SSIS SEL Screening SE Composite	SSIS SEL Screening AF Composite	SIBS/SEBS	SSIS-PSG Motivation to Learn	SSIS-PSG Prosocial Behavior
Number Identified At-Risk (N = 192)	67	44	63	32	24
Percentage Identified At-Risk	34.9%	22.9%	32.1%	16.7%	12.5%

Additionally, a ROC analysis was conducted using the SSIS SEL Screening AF Composite scale and the SSIS-PSG Reading and Mathematics domains as criterion. Figures 5 and 6 illustrate the AUC analysis SSIS SEL Screening AF Composite scale when predicting membership in the academically at-risk group. The AUC using the PSG Reading domain as criterion is .786, which is considered to be a good discrimination percentage (see Figure 5). For the comparison of the SSIS-PSG Reading and the SSIS SEL Screening AF Composite, the sensitivity index was .632, and the specificity index was .858. The positive predictive power was .818, and the negative predictive power was .941. The AUC using the PSG Mathematics domain as criterion is .771, which is considered to be a good discrimination percentage (see Figure 6). For the comparison of the SSIS-PSG Mathematics and the SSIS SEL Screening AF Composite, the sensitivity index was .638, and the specificity index was .903. The positive predictive power was .682, and the negative predictive power was .885. The SSIS SEL Screening AF Composite

conditional probabilities indices are presented in Table 13. Overall, the SSIS SEL Screening AF ROC analyses results and classification accuracy indices were consistent with previous findings (Elliott et al., in press).

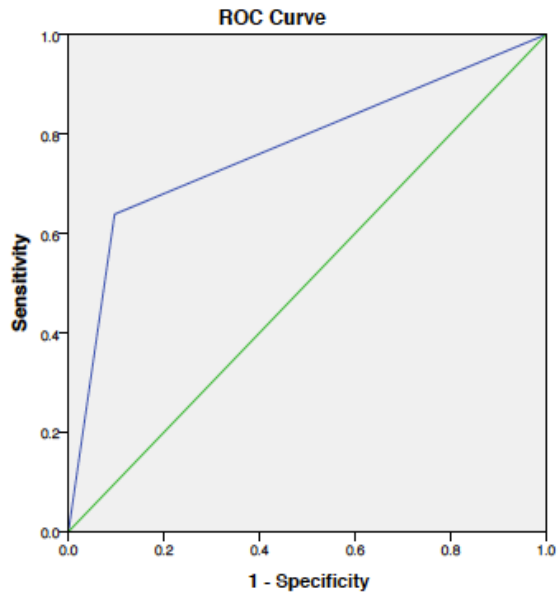


Figure 5. ROC AUC for the SSIS SEL Screening AF Composite and SSIS-PSG Reading

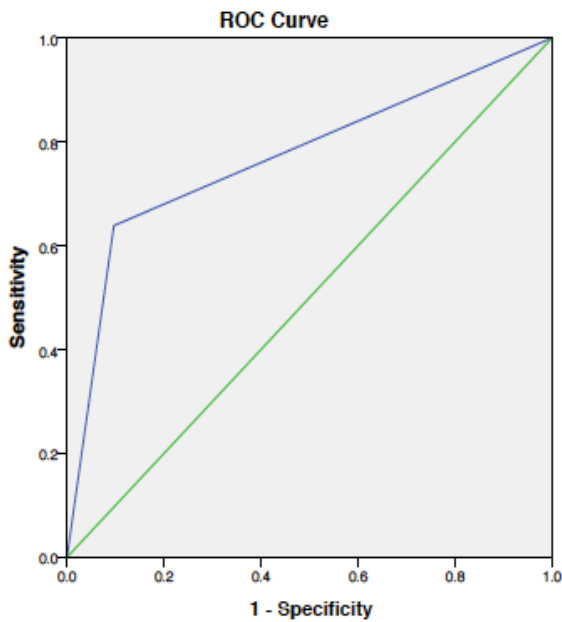


Figure 6. ROC AUC for the SSIS SEL Screening AF Composite and SSIS-PSG Mathematics

Table 13  
Summary of Classification Accuracy of SSIS SEL Screening AF Composite by SSIS-PSG  
Criterion Measure

	PSG Reading	PSG Mathematics
Sensitivity	63.2%	63.8%
Specificity	85.8%	90.3%
Positive Predictive Power	81.8%	68.2%
Negative Predictive Power	94.1%	88.5%

### Predictive Validity

To evaluate short-term predictive validity, the association between the SSIS SEL Screening/Progress Monitoring Scales and student academic achievement and behavioral data was assessed. Data from the Time 1 administration of the SSIS SEL Screening and other social emotional and behavior screeners collected approximately 7 weeks into the academic calendar was used, as well as student-level data collected from teachers at the end of the first academic semester. Pearson's product-moment correlation coefficients were calculated to assess the strength of the relationship between screener scores and school performance indicators. As noted previously, the strength of the resulting correlations was interpreted using the following recommendation: less than .30 are small, .30 to .49 are moderate, and greater than .50 are strong (Cohen, 1977). Results are located in Table 14.

The SSIS SEL Screening SE Composite was moderately and negatively correlated with the number of reported ODRs ( $r = -.36$ ) and suspensions ( $r = -.39$ ), which demonstrates that the lower the teacher-perceived social emotional skills, the greater number of ODRs and suspensions

students tended to receive during the academic semester. A small negative correlation was also found between the SSIS SEL Screening AF Composite and reported suspensions ( $r = -.21$ ), indicating that as teacher-reported academic skills decreased, student suspensions increased. The SSIS SEL Screening SE and AF Composites also resulted in moderate ( $r = .47$  and  $r = .42$ ) and strong positive correlations ( $r = .58$  and  $r = .51$ ) with ELA and Math GPA, respectively. When considering academic performance, as teacher-reported social emotional and academic functioning scores increased, so did their reported ELA and Math academic achievement.

In addition, the relationship between the combined SIBS/SEBS and SSIS-PSG behavior domain ratings and school performance indicators was investigated. These results are also summarized in Table 14. The combined SIBS/SEBS was moderately correlated with both teacher-reported ODRs ( $r = .37$ ) and suspensions ( $r = .33$ ), indicating that as teacher-perceived internalizing and externalizing behavior increased, the greater number of reported ODRs and suspensions students tended to earn. A small correlation was found between the SIBS/SEBS and teacher-reported absences ( $r = .14$ ). The combined SIBS/SEBS was also moderately and negatively correlated with ELA GPA ( $r = -.32$ ). A small negative correlation was revealed between the combined SIBS/SEBS and teacher-reported Math GPA ( $r = -.20$ ). Overall, when considering academic performance, greater emotional and behavioral problems as reported on the SIBS/SEBS were associated with poorer academic grades.

A moderately-strong relationship was also revealed between higher ratings on the SSIS-PSG Prosocial Behavior domain and fewer ODRs earned ( $r = -.32$ ), and a small negative correlation was found between higher ratings in the domain and suspensions ( $r = -.22$ ). A small relationship was also shown between lower ratings on the SSIS-PSG Motivation to Learn domain and greater attendance ( $r = -.22$ ). Lastly, the SSIS-PSG Motivation to Learn and



Prosocial Behavior domains resulted in moderate correlations with teacher reported ELA ( $r = .44$  and  $r = .42$ , respectively) and Math GPA ( $r = .49$  and  $r = .42$ , respectively), indicating that as teacher-reported motivation to learn and social skills increased, students tended to earn higher academic grades.

Table 14  
Correlations Between Screener Scores and School Performance Indicators

	ODRs	Suspensions	Absences	ELA GPA	Math GPA
SEL Screening SE Composite	-.36*	-.39*	-.08	.47*	.42*
SEL Screening AF Composite	-.11	-.21*	-.12	.58*	.51*
SIBS/SEBS	.37*	.33*	.14*	-.32*	-.20*
SSIS-PSG Motivation to Learn	-.17	-.16	-.22*	.44*	.49*
SSIS-PSG Prosocial Behavior	-.32*	-.22*	-.11	.42*	.42*

\*Pearson correlation is significant,  $p < .01$ .

### Social Validity

Teacher ratings on the Usage Rating Profile – Assessment (URP-A; Chafouleas et al., 2012) regarding the social validity and usability of the SSIS SEL Screening/Progress Monitoring Scales were analyzed using the mean item scores on the Acceptability and Feasibility subscales and Total Usability score. The average ratings and standard deviations on the URP-A are presented in Table 15, with higher mean item scores suggestive of greater perceived acceptability, feasibility, and usability of the SSIS SEL Screening/Progress Monitoring Scales as rated by teachers. Teachers' mean item ratings on the URP-A regarding the SSIS SEL Screening measure fell in the *Agree* range for the Acceptability ( $M = 4.48$ ,  $SD = 0.80$ ) and Feasibility ( $M =$

4.53,  $SD = 0.62$ ) subscales, as well as the Total Usability scale ( $M = 4.50$ ,  $SD = 0.75$ ). Overall, this suggests that teachers perceived the SSIS SEL Screening measure as being largely acceptable, feasible, and usable.

Table 15  
Teacher-Reported Acceptability and Usability of the SSIS SEL Screening

URP-A Results from Teachers	<i>M</i>	<i>SD</i>
Total Usability	4.50	0.75
Acceptability	4.48	0.80
Feasibility	4.53	0.62

## CHAPTER 4 DISCUSSION

Social emotional skills are integral to student success at school, home, and in the larger community. Social skills enable students to navigate interpersonal domains and are involved in the development and maintenance of relationships. Social, emotional, and behavioral concerns range on a continuum, and extant research consistently demonstrating that deficits are associated with poorer outcomes in and outside of school. As such, many schools have adopted universal screening methods to prevent or moderate adverse outcomes and promote school and student success. The purpose of the current study was to extend existing research by Elliott et al. (in press), and further investigate the technical adequacy, classification accuracy, and usability of the SSIS SEL Screening/Progress Monitoring Scales. Specifically, the research questions examined in this study evaluated the internal consistency and stability of SSIS SEL Screening/Progress Monitoring Scales scores, the association between the SSIS SEL Screening/Progress Monitoring Scales scores and scores on other well-established screening measures and academic and behavioral school performance indicators, rates of risk detection and classification accuracy, and teacher-perceived acceptability, feasibility, and usability of the measure.

The first research question examined the adequacy of evidence for the internal consistency and test-retest reliability of the SSIS SEL Screening/Progress Monitoring Scales. Teachers' ratings on the SSIS SEL Screening/Progress Monitoring Scales resulted in internal consistency reliability estimates of at least .70 at both Time 1 and Time 2 administration (Cronbach's  $\alpha = .93-.96$ ), indicating adequate internal consistency reliability for the overall screening measure. The SE and AF Composites also yielded adequate internal consistency estimates at both administrations (SE  $\alpha = .93-.96$ ; AF  $\alpha = .87-.91$ ), suggesting adequate internal

consistency reliability in comparison to the recommended reliability coefficient cutoff for both the SE and AF Composite scales. These findings are consistent with previous research on the SSIS SEL Screening/Progress Monitoring Scales that found adequate internal consistency reliability coefficients ranging from .90 to .93 (Elliott et al., in press). The stability of scores or test-retest reliability was also computed, and reliability coefficients were considered adequate if .70 or higher. Teacher ratings on the SSIS SEL Screening/Progress Monitoring Scales SE Composite scale resulted in a reliability coefficient of .68, with reliability of subscales ranging from .61 to .68. Additionally, the AF Composite scale yielded a test-retest reliability coefficient of .58, with reliability of subscales ranging from .56 to .64. Although approaching the threshold, test-retest reliability estimates were below the .70 cutoff suggested for screening measures. Overall, the internal consistency and test-retest reliability estimates for the AF Composite scale were slightly lower than the SE Composite scale. Extending the findings of Elliott et al. (in press), the current study demonstrates that the SSIS SEL Screening SE Composite has the potential to yield reliable and precise scores. The current study also provides new evidence for the stability of SE subscales scores, indicating that the SSIS SEL Screening SE Composite scales generate scores that are approaching the threshold of .70.

The second research question investigated the association of scores on the SSIS SEL Screening/Progress Monitoring Scales and other well-established universal screening measures, including the SSIS-PSG and the combined SIBS/SEBS. Concurrent validity estimates were computed using Time 1 administration scores. All correlations between the screening measures and the SSIS SEL Screening/Progress Monitoring Scales SE Composite were within the large range ( $>.50$ ; Cohen, 1988), including the SSIS-PSG, combined SIBS/SEBS, and SSIS SEL Core Skills Scale, and SSIS SEL Screening AF Composite scale. The direction of the associations

between scores on the SSIS SEL Screening SE Composite and other screening measures were consistent with the types of behavior assessed. Specifically, the SSIS SEL Screening SE Composite yielded strong and significant positive correlations with the SSIS SEL Screening AF Composite, the SSIS-PSG Motivation to Learn and Prosocial Behavior domains, and the SSIS SEL Core Skills Scale. As total scores on the SSIS SEL SE Composite increased, teacher-reported greater social emotional and academic skills, including overall academic functioning and academic engagement in the classroom. Additionally, the SSIS SEL Screening SE Composite yielded a strong and significant negative correlation with the combined SIBS/SEBS global rating, indicating that as social emotional skills increased, teachers reported greater number and/or frequency of problem behaviors. Importantly, the largest correlation with the SSIS SEL Screening SE Composite was found with the combined SIBS/SEBS ( $r = -.68$ ), which served as the initial criterion measure in the current study. The SSIS SEL AF Composite scale also resulted in significant and strong correlations with the SSIS-PSG, indicating that increased total scores on the AF Composite were related to greater prosocial behaviors and academic engagement. The SSIS SEL AF Composite scale yielded a small negative association with the combined SIBS/SEBS. The SSIS SEL AF Composite scale assesses academic engagement and reading and math skills, while the combined SIBS/SEBS assesses a variety of internalizing and externalizing behaviors. The differences in behaviors assessed by the two measures may account for the weaker association between the scores. Overall, the strength of the associations between the SSIS SEL Screening/Progress Monitoring Scales and the SSIS-PSG and combined SIBS/SEBS provides further support for the concurrent validity of the SE and AF Composite scales.

Third, the rate of risk detection was examined across all universal social emotional and behavioral screening measures. Using the previously defined cutoff score (Elliott et al., in press), the SSIS SEL Screening SE Composite classified 10.9% of the students in the current sample as at-risk. The rate of risk detection by the SSIS SEL Screening SE Composite scale was then compared to the criterion universal screening measures to examine whether the rate of detection was similar. The combined SIBS/SEBS identified 32.1% of the sample as at-risk, and the SSIS-SG classified 12.5% on the Prosocial Behavior domain and 16.7% on the Motivation to Learn domain. The SSIS SEL Screening SE Composite and SSIS-PSG Prosocial Behavior domain were more similar, likely due to the consistency in behaviors assessed. However, the combined SIBS/SEBS resulted in 42 additional students identified as at-risk as compared to the SSIS SEL Screening SE Composite. To further analyze risk detection among the current sample of students, the rate of risk identification was also examined for the SSIS SEL Screening SE scales, including the Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision-Making scales. Teacher ratings on the Self-Management scale yielded the greatest number of at-risk students (22.4%), followed by the Social Awareness (19.8%) and Self-Awareness (19.3%) scale. Overall, ratings on the five SE scales resulted in greater number and percentage of students identified as at-risk as compared to the SSIS SEL Screening SE Composite scale.

Risk classification according to the SSIS SEL Screening/Progress Monitoring Scales SE and AF Composite scales is largely consistent with proposed percentages of at-risk students according to a MTSS model. In a MTSS model, approximately 15% of students do not adequately respond to universal core instruction and would benefit from additional intervention supports at the Tier 2 and 3 levels. In the present study, universal behavior and social emotional

screener identification rates ranged from 10.9% according to the SSIS SEL Screening SE Composite to 32.1% according to the combined SIBS/SEBS. Additionally, the updated cut score of 15 increased the percentage of identified students according to the SSIS SEL Screening SE Composite from 10.9% to 34.9%, which is more consistent with the number of students identified by the combined SIBS/SEBS and the individual SE scales. The SSIS SEL Screening SE scales (i.e., Self Awareness, Social Awareness, Responsible Decision-Making, Self Management, Relationship Skills) particularly mirror the MTSS proposed percentages, with 14.6 to 22.4 percent of students identified as at-risk according to the five social emotional scales.

The accuracy of risk detection was also examined using the combined SIBS/SEBS and SSIS-PSG as criterion. Accuracy in risk detection is an important consideration, as resources in schools are often limited and conditional probability indices provide further confidence that students are being accurately identified and resources are being allocated appropriately. Sensitivity, specificity, positive predictive power, and negative predictive power of the SSIS SEL Screening SE Composite were initially calculated using the combined SIBS/SEBS as criterion. The sensitivity of the SSIS SEL Screening SE Composite (.286) was well below the recommended cutoff of .60, indicating that the SSIS SEL Screening SE Composite missed the identification of approximately 71.4% of students identified by the combined SIBS/SEBS as at-risk. The specificity index indicated that classification of students not at-risk was highly accurate, with 97.7% of students found not at-risk by the combined SIBS/SEBS were identified as non-risk by the SSIS SEL Screening SE Composite scale. The positive predictive power (85.7%) and negative predictive power (73.7%) of the SSIS SEL Screening SE Composite scale by the combined SIBS/SEBS criterion measure suggested adequate accuracy and low number of false positives and false negatives.

Given the greater consistency in at-risk identification rates between the SSIS SEL Screening SE Composite and SSIS-PSG Prosocial Behavior domain, conditional probability indices were also calculated using the SSIS-PSG Prosocial Behavior domain as criterion, and therefore replicating the Elliott et al. (in press) study conducted in Australia. As with the combined SIBS/SEBS as criterion, the SSIS SEL Screening SE Composite scores resulted in inadequate sensitivity (.417), indicating that the SSIS SEL Screening SE Composite missed the identification of approximately 58.3% of students identified by the combined SSIS-PSG Prosocial Behavior domain as at-risk. Positive predictive power was also inadequate, which suggests greater false positives identified by the SSIS SEL Screening SE Composite scale as compared to the SSIS-PSG Prosocial Behavior domain. The resulting specificity (.935) and negative predictive power (.91) exceeded the recommended cutoff of 0.60. Overall, results suggest that the SSIS SEL Screening SE Composite does not adequately identify students as at-risk as compared to other well-research social emotional and behavioral screeners. As such, ROC analyses were conducted to determine the best cutoff score for the SSIS SEL Screening/Progress Monitoring Scales SE Composite to maximize sensitivity and specificity.

ROC curve analyses were conducted using both the combined SIBS/SEBS and the SSIS-PSG Prosocial Behavior domain as criterion. The analyses revealed that a score of 15 was the optimal cutoff point of hit and false positive rates. The area under the curve (AUC) was .84 and .87 using the SIBS/SEBS and SSIS-PSG, respectively, which are considered to be good discrimination percentages. Following the ROC analyses, conditional probabilities were again computed using the adjusted cutoff score of 15 and the combined SIBS/SEBS and SSIS-PSG as criterion. The sensitivity of the SSIS SEL Screening SE Composite was adequate (.683), indicating that students are identified by the screening measure as at-risk with sufficient accuracy



according to the combined SIBS/SEBS criterion measure. Although the sensitivity index met the cutoff of .60 with the adjusted cut score, the results suggest that the SSIS SEL Screening SE Composite missed the identification of 31.7% of students identified by the SIBS/SEBS as at-risk. However, the sensitivity index was much higher when using the SSIS-PSG Prosocial Behavior domain as criterion (91.7%), indicating that the SSIS SEL Screening SE Composite more accurately identifies students as at-risk according to the SSIS-PSG criterion. The specificity index was adequate when using both the combined SIBS/SEBS and SSIS-PSG as criterion, 81.4% and 73.2%, respectively.

Positive predictive power was adequate using the combined SIBS/SEBS (.642), but missed the recommended cutoff for classification accuracy when using the SSIS-PSG as criterion (.328). Low positive predictive power in the current study indicated that the SSIS SEL Screening SE Composite identified additional students not identified by the criterion measure SSIS-PSG, resulting in false positives or over-identification of risk according to the SSIS-PSG. However, this is generally acceptable for screening purposes, particularly if schools utilize multi-gated screening methods. Negative predictive power of the SSIS SEL Screening was adequate for both the combined SIBS/SEBS criterion (.84) and the SSIS-PSG Prosocial Behavior domain criterion (.984), indicating that the SSIS SEL Screening SE Composite identified a low number of false negatives.

In addition, ROC analyses and conditional probabilities were also computed for the SSIS SEL Screening AF Composite scale using the SSIS-PSG Reading and Mathematics domain as criterion, as examined by Elliott et al. (in press). The area under the curve (AUC) was .786 and .771 using the PSG Reading and Mathematics domain, respectively, which are considered to be good discrimination percentages. Overall, classification accuracy indices obtained were

consistent with previous findings, indicating that the AF Composite adequately classifies students as at-risk according to the SSIS-PSG Reading and Mathematics domain.

In regards to reported conditional probabilities indices, it is important to note that not all indices can be simultaneously increased. Therefore, schools should evaluate the reported conditional probabilities in relation to severity of problem behaviors and associated outcomes and the availability of resources in schools, such as time, staff, and intervention materials. In addition, future research should examine the updated cutoff score, particularly with a large and more variable sample of students. Until further research is conducted, implementation in schools using the updated cut score of 15 should be interpreted cautiously, as conditional probability indices indicate possible over-identification and false positives in comparison to the SSIS-PSG Prosocial Behavior domain. However, schools may utilize multi-gate screening methods to ensure accurate identification and thereby decrease the unnecessary use of resources for students not at-risk for social emotional and behavioral problems. It may also be of use to assess domain score elevations (i.e., Self Awareness, Social Awareness, Responsible Decision-Making, Self Management, Relationship Skills) and target deficits accordingly. For example, school administrators and decision making teams could examine scores of all students identified as at-risk according to the SSIS SEL Screening SE Composite, and subsequently assign students into groups by the individual domain deficits.

The fourth research question explored the relationship with academic and behavioral school performance indicators and scores on the SSIS SEL Screening/Progress Monitoring Scales and other universal behavior screeners. Regarding behavioral school performance indicators, significant correlations were found between office discipline referrals (ODRs) and scores on the SSIS SEL Screening SE Composite, combined SIBS/SEBS, and the SSIS-PSG

Prosocial Behavior domain. The strength of the correlations between ODRs and the universal screening measures were moderate, and in the expected direction. In general, the greater number of teacher-reported ODRs earned by students, the greater number and/or severity of reported problem internalizing and externalizing behaviors and poorer social emotional skills. The relationship between ODRs and SSIS SEL Screening SE Composite, combined SIBS/SEBS, and SSIS-PSG Prosocial Behavior domain scores were similar in strength, providing comparable evidence for the predictive validity of the measures. The relationship between screener scores and teacher-reported suspensions were also examined, resulting in small correlations between suspensions and the SSIS SEL Screening AF Composite and SSIS-PSG Prosocial Behavior domain scores, and moderate correlations between suspensions and the SSIS SEL Screening SE Composite and combined SIBS/SEBS scores. All correlations were in the expected directions. As with ODRs, students with lower the teacher-perceived social emotional skills and greater reported problem behaviors tended to receive more suspensions during the academic semester. Additionally, when comparing the SSIS SEL Screening/Progress Monitoring scales to the criterion screening measures, the significant relationships with suspensions were largely similar in strength as compared to the combined SIBS/SEBS and SSIS-PSG Prosocial Behavior domain. Overall, ODRs and suspensions and universal behavior screener score correlations were small-to-moderate. However, it is likely that the range of behaviors assessed (i.e., both internalizing and externalizing behaviors and variety of social emotional skill domains) may account for the small-to-moderate strength of the correlations. Additionally, only a small percentage of students earned ODRs (16% of the overall sample) and suspensions (4% of the overall sample), which may also have affected the relationship between reported ODRs and suspensions and screening measure scores. Lastly, teacher-reported absences revealed small correlations with the combined

SIBS/SEBS and SSIS-PSG Motivation to Learn domain, both in the expected direction. This is largely consistent with previous research examining universal behavior screener scores and student attendance.

Regarding academic school performance indicators, the evidence for an association between SSIS SEL Screening SE and AF Composite scores and ELA and Math GPA was moderate to strong ( $r = .42$  to  $.58$ ). As expected, students with greater social emotional skills and academic engagement tended to earn higher average grades in Math and ELA. Although the strength of the association was stronger for the AF than SE Composite scale, the moderate association between teacher-reported social emotional skills and academic performance is informative. Additionally, the SSIS-PSG Motivation to Learn and Prosocial Behavior domains also yielded moderate correlations with ELA and Math GPA ( $r = .42$  to  $.49$ ), with greater academic engagement and prosocial skills being associated with higher academic grades. Additionally, the evidence for an association between the combined SIBS/SEBS and ELA and Math GPA was small-to-moderate ( $r = -.20$  to  $-.32$ ), and in the expected direction. Generally, as student social emotional and behavioral problems increased, math and ELA grades decreased. These findings are consistent with extant research describing social emotional skills as academic enablers and problem behaviors as academic disablers (Gresham & Elliott, 2014). Specifically, students with social emotional skills deficits are at risk for a host of negative academic outcomes (e.g., poor academic performance, school dropout), and short-term predictive data obtained in the present study further support these findings. However, it would be beneficial for future studies to further examine these relationships.

The final research question examined teacher perceptions of the acceptability, feasibility, and usability of the SSIS SEL Screening/Progress Monitoring Scales. Acceptability is defined as

the perceived appropriateness of the assessment and interest in its use. Teacher ratings on the URP-A revealed a mean item score of 4.48 on the Acceptability subscale, indicating that teachers in the current sample find the screening measure to be acceptable. Feasibility concerns the perceived ease of use, particularly in terms of cost of the measure, time and effort to complete, and fit with the school-based services. Teacher ratings revealed a mean score of 4.53 on the Feasibility subscale, indicating that teachers in the current sample rate the SSIS SEL Screening/Progress Monitoring Scales as largely easy to use. Overall, perceived usability revealed a mean score of 4.50, indicating that teachers found the SSIS SEL Screening/Progress Monitoring Scales to be easy to use, time efficient, acceptable, and relevant to efforts to improve student social emotional skills. These indicators are important considerations because poor usability, feasibility, and acceptability are likely to serve as a barrier to screening implementation and data collection in schools. Specifically, if teachers perceive screener administration as too time intensive and requiring an excessive amount of effort, they are likely to be more resistant and/or haphazardly answer screener questions.

### **Limitations and Future Directions**

Although the current study provides evidence as to the usefulness of the SSIS SEL Screening/Progress Monitoring Scales for use in schools, there remain several limitations and avenues for future directions in research. Data examined in the present study was exclusively obtained via teacher report. Due to practicality concerns, parent consent and child assent was not obtained. As an alternative, teachers utilized student codes when reporting student ODRs, suspensions, absences, and academic performance, which allowed the research to bypass collecting identifiable student data. As indicated previously, biases and recollection issues may have affected teachers' report of student behavioral and academic performance indicators.

However, despite the lack of individual student-level data obtained in this study, bypassing the active parent consent requirement addresses some of the limitations inherent in social emotional and behavior screener research. Specifically, collecting teacher-report social emotional skills and behavioral indicators at the classwide level increases the variability of data obtained. Obtaining data from all students in each classroom is also more applicable to actual screener implementation and use in the school, as teachers would typically rate all students in their respective classrooms. However, perhaps a more informative metric in the future would be to obtain individual teacher and aggregate school level data, as well as informal “time-aways” from the classroom and objective behavior indicators, such as behavior observations in and outside of the classroom.

Further, the present study utilized a single screening methodology, with no follow-up systematic interventions administered. The implementation of the SSIS SEL Screening/Progress Monitoring Scales and the corresponding SSIS SEL Classwide Intervention Program would be potential direction for research to explore. Students identified as at-risk according to the SSIS SEL Screening/Progress Monitoring Scales may be placed in a systematic intervention targeting reported skill deficits. The SSIS SEL Screening/Progress Monitoring Scales may then be used to progress monitor student response to intervention. Additionally, the screening measure may also be used as an indicator of intervention effectiveness. As Elliott et al. (in press) purported, the SSIS SEL Screening/Progress Monitoring Scales has the potential to be used as a progress monitoring tool in schools. However, generalized use for progress monitoring requires that additional studies examine the repeated use of the measure with students in and out of related interventions.

Invariance across gender and race/ethnicity was also not examined in the present study. This area of research is important, and future studies should examine possible differences in teacher reporting and student data across groups. Additionally, generalizability of the current findings is limited by the sample obtained in the present study. In future studies, it would also be useful to obtain more ethnically diverse samples and greater representation of preschool, elementary and middle school grades.

Despite the limitations, findings from the current study provide important implications for research and practice in universal screening, identification, and intervention implementation in schools. The present study extends research on the SSIS SEL Screening/Progress Monitoring Scales and provides evidence for the technical adequacy, classification accuracy and usability in schools. The SSIS SEL Screening/Progress Monitoring Scales demonstrated adequate reliability, concurrent validity with other well-established social skills and behavior screening measures, and short-term predictive validity related to student academic performance and behavioral problems. Findings also support previous research on the classification accuracy and identification rates of at-risk students. The rates of risk detection were largely consistent with those identified at at-risk and not at-risk by the SSIS-PSG and combined SIBS/SEBS. Lastly, teachers in the present study generally rated the SSIS SEL Screening/Progress Monitoring Scales are acceptable, feasible and useable.

## REFERENCES

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- Agostin, T. M., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools, 34*(3), 219-228.
- 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.
- Ashdown, D. M., & Bernard, M. E. (2012). Can explicit instruction in social and emotional learning skills benefit the social-emotional development, well-being, and academic achievement of young children? *Early Childhood Education Journal, 39*, 397-405. doi:10.1007/s10643-011-0481-x
- Barkley, R. (2010). Deficient emotional self-regulation: A core component of attention-deficit/hyperactivity disorder. *Journal of ADHD and Related Disorders, 1*, 5-37.
- Benson, P. L. (2006). *All kids are our kids: What communities must do to raise responsible and caring children and adolescents*. San Francisco: Jossey-Bass.
- Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science, 11*, 302-306.
- Collaborative for Academic, Social, and Emotional Learning. (2012). *2013 CASEL guide: Effective social and emotional learning programs—Preschool and elementary school edition*. Chicago, IL: Author.
- Cook, C. R., Frye, M., Slemrod, T., Lyon, A. R., Renshaw, T., & Zhang, Y. An integrated approach to universal prevention: Independent and combined effects of PBIS and SEL on youths' mental health. *School Psychology Quarterly, 30*(2), 166-183.
- Cook, C. R., Volpe, R. J., & Gresham, F. M. (2012). *Technical adequacy, classification accuracy and social validity of the Student Externalizing Behavior Screener*. Unpublished manuscript.



- Costello, J. E., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, *60*, 837–844.
- Crowe, L. M., Beauchamp, M. H., Catroppa, C., & Anderson, V. (2011). Social function assessment tools for children and adolescents: A systematic review from 1988 to 2010. *Clinical Psychology Review*, *31*, 767-785. doi:10.1016/j.cpr.2011.03.008
- Denham, S. A. (2015). Assessment of SEL in educational contexts. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.) *Handbook of social and emotional learning: Research and practice* (pp. 3-19). New York, NY: The Guilford Press.
- Denham, S. A., Bassett, H. H., Zinsser, K., & Wyatt, T. M. (2014). How preschoolers' social-emotional learning predicts their early school success: Developing theory-promoting, competency-based assessments. *Infant and Child Development*, *23*, 426-454. doi:10.1002/icd.1840
- 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. doi: 10.1177/1063426611417627
- Dowdy, E., Kamphaus, R., Twyford, J., & Dever, B. (2014). Culturally competent emotional and behavioral screening. In M. Weist, N. Lever, C. Bradshaw, & J. Owens (Eds.), *Handbook of school mental health*, (pp. 311-322). New York, NY: Springer.
- Drummond, T. (1994). *The Student Risk Screening Scale (SRSS)*. Grants Pass, OR: Josephine County Mental Health Program.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405-432. doi:10.1111/j.1467-8624.2010.01564.x
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, *45*, 294-309. doi:10.1007/s10464-010-9300-6
- Elias, M. J., & Weissberg, R. P. (2000). Primary prevention: Educational approaches to enhance social and emotional learning. *Journal of School Health*, *70*(5), 186-190.
- Elliot, S. N., Davis M. D., Frey, J. R., & Cooper, G. (2017) The development and initial validation of a social emotional learning assessment for universal screening. *Journal of Applied Development Psychology*, in press.

- Frey, A. J., Lingo, A., & Nelson, C. M. (2010). Implementing positive behavior support in elementary schools. In M. Shinn & H. Walker (Eds.) *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 293-312). Bethesda, MD: National Association of School Psychologists.
- Gerber, M. M., & Semmel, M. I. (1984). Teacher as imperfect test: Reconceptualizing the referral process. *Educational Psychologist, 16*, 137-148. doi: 10.1177/0022219409335217
- Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135. doi: 10.1016/j.jsp.2006.05.005
- Graczyk, P. A., Matjasko, J. L., Weissberg, R. P., Greenberg, M. T., Elias, M. J., & Zins, J. E. (2000). The role of the collaborative to advance social and emotional learning (CASEL) in supporting the implementation of quality school-based prevention programs. *Journal of Educational and Psychological Consultation, 11*(1), 3-6.
- Graziano, P. A., & Hart, K. (2016). Beyond behavior modification. Benefits of social-emotional/self-regulation training for preschoolers with behavior problems. *Journal of School Psychology, 58*, 91-111. <http://dx.doi.org/10.1016/j.jsp.2016.07.004>
- Graziano, P. R. & Hart, K. (2016). Beyond behavior modification: Benefits of social-emotional/self-regulation training for preschoolers with behavioral problems. *Journal of School Psychology, 58*, 91-11. <http://dx.doi.org/10.1016/j.jsp.2016.07.004>
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*(6/7), 466-474. doi:10.1037/0003-066X.58.6-7.466
- Gresham, F. M. (2005). Response to intervention: An alternative means of identifying students as emotionally disturbed. *Education and Treatment of Children, 28*, 328-344.
- Gresham, F. M. (2010). Evidence-based social skills interventions: Empirical foundations for instructional approaches. In M. Shinn & H. Walker (Eds.) *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 337-362). Bethesda, MD: National Association of School Psychologists.
- Gresham, F. M., & Elliot, S. N. (2014). Social skills assessment and training in emotional and behavioral disorders. In H. M. Walker & F. M. Gresham (Eds.) *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools*. (pp. 152-172). New York, NY: The Guilford Press.
- Gresham, F. M., Van, M. B., & Cook, C. R. (200). Social skills training for teaching replacement behaviors: Remediation acquisition deficits in at-risk students. *Behavioral Disorders, 31*(4), 363-377.

- Halle, T. G., & Darling-Churchill, K. E. (2016). Review of measures of social and emotional development. *Journal of Applied Developmental Psychology, 45*, 8-18. <http://dx.doi.org/10.1016/j.appdev.2016.02.003>
- Hartman, K., Gresham, F., & Byrd, S. (2017). Student internalizing and externalizing behavior screeners: Evidence for reliability, validity, and usability in elementary schools. *Behavior Disorders, 42* (3), 108-118.
- Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 *et seq.* (2004) (reauthorization of Individuals with Disabilities Act of 1990)
- Jones, D., Greenberg, M. T., & Crowley, M. (2015). The economic case for SEL. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.) *Handbook of social and emotional learning: Research and practice* (pp. 213-228). New York, NY: The Guilford Press.
- Jones, S. M., & Bouffard, S. M. (2012). Social and emotional learning in schools from programs to strategies. *Social Policy Report, 26*(4), Society for Research in Child Development.
- Kamphaus, R. W., & Reynolds, C. R. (2007). *BASC-2 Behavioral and Emotional Screening System*. Minneapolis, MN: Pearson Assessment.
- Kettler, R. J., Glover, T. A., Albers, C. A., & Feeney-Kettler, K. A. (2014). An introduction to universal screening in educational settings. In R. J. Kettler, T. A. Glover, C. A. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational settings: Evidence-based decision making for schools* (pp. 249-273). Washington, DC: American Psychological Association.
- Krull, J., Wilbert, J., & Hennemann, T. (2014). The social and emotional situation of first graders with classroom behavior problems and classroom learning difficulties in inclusive classes. *Learning Disabilities: A Contemporary Journal, 12*(2), 169-190.
- Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45*, 163-191. doi:10.1016/j.jsp.2006.11.005
- Maag, J.W. (2006). Social skills training for students with emotional and behavioral disorders: A review of reviews. *Behavioral Disorders, 32* (1), 5-17.
- Malecki, C.K., & Elliot, S. B. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*, 1-23.
- McLeod, J. D., & Fettes, D. L. (2007). Trajectories of Failure: The Educational Careers of Children with Mental Health Problems. *AJS; American Journal of Sociology, 113*(3), 653-701. <http://doi.org/10.1086/521849>

- Merrell, K. W., & Walker, H. M. (2004). Deconstructing a definition: Social maladjustment versus emotional disturbance and moving the EBD field forward. *Psychology in the Schools, 41*(8), 899-910. doi: 10.1002/pits.20046
- McBride, A. M., Chung, S., & Robertson, A. (2016). Preventing academic disengagement through a middle school-based social and emotional learning program. *Journal of Experiential Education, 39*(4), 370-385. doi:10.1177/1053825916668901
- National Association of School Psychologists (2002). *Position statement on students with emotional and behavioral disorders*. Bethesda, MD: Author
- National Commission on Excellence in Education. (1983). *Report of the National Commission on Excellence in Education*. Washington, DC: Author.
- Payton, J. W., Wardlaw, D. M., Graczyk, P.A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. Social and emotional learning: A framework for promoting mental health and reducing risk behavior in children and youth. *Journal of School Health, 70*(5), 179-185.
- Rabiner, D. L., Godwin, J., & Dodge, K. A. (2016). Predictive academic achievement and attainment: The contribution of early academic skills, attention difficulties, and social competence. *School Psychology Review, 45*(2), 250-267
- Rutherford, R. B. Jr., Quinn, M. M., & Mathur, S.R. (2004). *Handbook of Research in Emotional and Behavioral Disorders*. New York, NY: The Guilford Press.
- Spielberger, C. D., Edwards, C.D., Lushene, R. E., Montuori, J., & Platzek D. (1973). *STAIC Preliminary Manual*. Palo Alto, CA: Consulting Psychologists Press.
- Squires, J. (2010). Designing and implementing effective preschool programs: A linked systems approach for social-emotional early learning. In M. Shinn & H. Walker (Eds.) *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 293-312). Bethesda, MD: National Association of School Psychologists.
- Stoiber, K. C. (2014). A comprehensive framework for multitiered systems of support in school psychology. In P. L. Harrison, & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 41-70). Bethesda, MD: National Association of School Psychologists.
- Strein, W., Hoagwood, K., & Cohn, A. (2003). School psychology: A public health perspective I. Prevention, populations, and systems change. *Journal of School Psychology, 41*, 23-38. [http://dx.doi.org/10.1016/S00224405\(02\)00142-5](http://dx.doi.org/10.1016/S00224405(02)00142-5)
- 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*, 79-96.

Walker, H. M., & Severson, H. (1992). *Systematic Screening for Behavior Disorders* (2nd ed.). Longmont, CO: Sopris West.

Walker, H. M., Severson, H. H., & Seeley, J. R. (2010). Universal, school-based screening for the early detection of behavioral problems contributing to later destructive outcomes. In M. R. Shinn, & H. M. Walker (Eds.), *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 677-702). Bethesda, MD: National Association of School Psychologists.

Weissberg, R. P. & Cascarino, J (2013). Academic learning + social-emotional learning = national priority. *Phi Delta Kappan, 95*(2): 8-13.

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullotta, T. P. (2015). Social emotional learning: Past, present, and future. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.) *Handbook of social and emotional learning: Research and practice* (pp. 3-19). New York, NY: The Guilford Press.

Wentzel, K. R. (1991). Social competence at school: Relation between social responsibility and academic achievement. *Review of Educational Research, 61*(1), 1-24.

Wiley, A. L., & Siperstein, G. N. (2015) SEL for students with high-incidence disabilities. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.) *Handbook of social and emotional learning: Research and practice* (pp. 213-228). New York, NY: The Guilford Press.

**APPENDIX A**  
**TEACHER DEMOGRAPHIC FORM**

Teacher Demographic Information

Teacher ID: \_\_\_\_\_

Grade Taught: \_\_\_\_\_

Number of Years in Education Field: \_\_\_\_\_

Age: \_\_\_\_\_

Sex (choose one):     Male    Female

Primary Ethnic Identity (choose one):

African American

Asian American

White, Non-Hispanic

Hispanic or Latino

Native American

Other (please specify): \_\_\_\_\_

How can we reach you most easily? (Please provide email and/or phone # [text]):

\_\_\_\_\_



## APPENDIX C SIBS AND SEBS COMBINED MEASURE

SIBS & SEBS

### Student Behavior Screening — Teacher Report

Student ID#: \_\_\_\_\_ Gender: MALE or FEMALE (circle one)

Teacher: \_\_\_\_\_ Grade: \_\_\_\_\_

Date: \_\_\_\_\_

Directions: For each item, please circle the response that best describes the student's behavior at school.

	Never	Rarely/ Seldom	Occasionally/ Moderately	Frequently/ Almost Always
1. Defiant or oppositional to adults.	0	1	2	3
2. Fights or argues with peers.	0	1	2	3
3. Bullies others.	0	1	2	3
4. Gets angry easily.	0	1	2	3
5. Lies to get out of trouble.	0	1	2	3
6. Disrupts class activities.	0	1	2	3
7. Has difficulty sitting still.	0	1	2	3
8. Nervous, worried, or fearful.	0	1	2	3
9. Bullied by peers.	0	1	2	3
10. Spends time alone.	0	1	2	3
11. Clings to adults.	0	1	2	3
12. Withdrawn.	0	1	2	3
13. Seems sad or unhappy.	0	1	2	3
14. Complains about being sick or hurt.	0	1	2	3



**APPENDIX D**  
**SSIS SEL EDITION – TEACHER FORM**  
**CORE SKILLS SCALE**

**SSIS SEL Edition – Teacher Form**  
**Core Skills**

**Date:**

**Student ID:**

**Instructions:** You will be presented with several statements that describe a student’s behavior. Please read each item and think about the student’s behavior during the past two months. Then, decide *how often* this student displays the behavior.

If the student *never* exhibits the behavior, select N.

If the student *seldom* exhibits the behavior, select S.

If the student *often* exhibits the behavior, select O.

If the student *almost always* exhibits the behavior, select A.

Remember: **N = Never, S = Seldom, O = Often, A = Almost Always**

- |                                                 |   |   |   |   |
|-------------------------------------------------|---|---|---|---|
| 1. Says “please.”                               | N | S | O | A |
| 2. Asks for help from adults.                   | N | S | O | A |
| 3. Follows your directions.                     | N | S | O | A |
| 4. Stays calm when teased.                      | N | S | O | A |
| 5. Acts responsibly when with others.           | N | S | O | A |
| 6. Pays attention to your instructions.         | N | S | O | A |
| 7. Takes care when using other people’s things. | N | S | O | A |
| 8. Stays calm when disagreeing with others.     | N | S | O | A |
| 9. Says “thank you.”                            | N | S | O | A |
| 10. Takes turn in conversations.                | N | S | O | A |

*\*This measure was adapted from the SSIS SEL Edition – Teacher Form (Gresham & Elliott, 2017)*

## APPENDIX E USAGE RATING PROFILE – ASSESSMENT



### Usage Rating Profile – Assessment

Name of Screener: \_\_\_\_\_

Date: \_\_\_\_\_

Teacher: \_\_\_\_\_

**Directions:** Consider the described assessment when answering each of the following statements. Circle the number that best reflects your agreement with the statement, using the scale provided below.

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
1. This assessment is an effective choice for understanding a variety of problems.	1	2	3	4	5	6
2. I would be able to allocate my time to implement this assessment.	1	2	3	4	5	6
3. The assessment is a fair way to evaluate the child's behavior problem.	1	2	3	4	5	6
4. The total time required to implement the assessment procedures would be manageable.	1	2	3	4	5	6
5. I would not be interested in implementing this assessment.	1	2	3	4	5	6
6. I would have positive attitudes about implementing this assessment.	1	2	3	4	5	6
7. This is a good way to assess the child's behavior problem.	1	2	3	4	5	6
8. Preparation of materials needed for this assessment would be minimal.	1	2	3	4	5	6
9. Material resources needed for this assessment are reasonable.	1	2	3	4	5	6
10. I would implement this assessment with a good deal of enthusiasm.	1	2	3	4	5	6
11. This assessment is too complex to carry out accurately.	1	2	3	4	5	6
12. Use of this assessment would not be disruptive to students.	1	2	3	4	5	6
13. I would be committed to carrying out this assessment.	1	2	3	4	5	6
14. The assessment procedures easily fit in with my current practices.	1	2	3	4	5	6
15. The amount of time required for record keeping would be reasonable.	1	2	3	4	5	6

*Selected items from URP-A, created by Sandra M. Chafouleas, Faith G. Miller, Amy M. Briesch, Sabina Rak Neugebauer, & Chris Riley-Tillman. Copyright © 2012 by the University of Connecticut. All rights reserved.*



**APPENDIX G**  
**SCHOOL ADMINISTRATOR INFORMED CONSENT FORM**

**LOUISIANA STATE UNIVERSITY**

**SCHOOL ADMINISTRATOR INFORMED CONSENT FORM**

We are requesting your approval and support to conduct the study *An Examination of the Technical Adequacy, Classification Accuracy, and Usability of the SSIS SEL Edition Screening/Progress Monitoring Scales in Elementary School Populations* at your school. The following sections outline the details of the study.

**Purpose of the Study:** Early identification and intervention for students at-risk for social, emotional, and behavioral difficulties is key mitigating negative outcomes. Universal screening is a proactive method for detecting at-risk students, and involves brief assessment of all students. This study is being conducted to examine a recently developed universal social-emotional screener for use in schools. In addition, this study will provide valuable knowledge for the purpose of comparing and contrasting the accuracy and acceptability of multiple social-emotional and behavior screeners.

**Study Procedures:** We are requesting approval to conduct a study on universal social-emotional and behavior screeners at your school. With your support, we will request participation in the study from teachers at your school. The study will involve three data collection periods. Participating teachers will be asked to complete several brief screeners on all students in their classroom approximately 6 weeks into the school year. On these screeners, teachers will rate each student's social-emotional skills, behavior and academic performance at school. One month later, teachers will complete one of the screeners again and fill out a short questionnaire indicating their opinion of the screeners. At the end of the semester we will collect academic and behavioral information for all students. To link student information across data collection periods, teachers will be sent a spreadsheet with anonymous codes to be assigned to each of their students and used on the questionnaires in lieu of their names. Once IDs are assigned, the spreadsheets are to be password protected by the teachers and resent to the researcher to manage. The research will not have access to teacher passwords. Details on this procedure will be provided at a training prior to the first data collection. Completion of the training, rating scales, and additional student information will take approximately 6-7 hours of each teacher's time over the course of the semester. Frank Gresham, Ph.D. and Shelby Byrd, M.S. 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 social-emotional 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 schools. In addition, to show our appreciation for teachers' assistance, we will provide each participating teacher with a \$10 gift card. All participating teachers will also be entered into a raffle to win a \$100 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 or fatigued from completing ratings of students' behavior. However, teachers will be trained on rating procedures to minimize these risks. Teachers are free to complete questionnaires at their leisure, such as over the weekend, after or before school.

**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 penalty.

**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 Shelby Byrd at sbyrd5@lsu.edu, Monday-Friday 8:00 a.m. – 4:30 p.m. If you have any questions about participants' 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: \_\_\_\_\_

**APPENDIX H**  
**TEACHER INFORMED CONSENT FORM**

**LOUISIANA STATE UNIVERSITY**

**TEACHER INFORMED CONSENT FORM**

We are requesting participation and collaboration in the study *An Examination of the Technical Adequacy, Classification Accuracy, and Usability of the SSIS SEL Edition Screening/Progress Monitoring Scales in Elementary School Populations* at your school. The following sections outline the details of the study.

**Purpose of the Study:** Early identification and intervention for students at-risk for social, emotional, and behavioral difficulties is key mitigating negative outcomes. Universal screening is a proactive method for detecting at-risk students, and involves brief assessment of all students. This study is being conducted to examine a recently developed universal social-emotional screener for use in schools. In addition, this study will provide valuable knowledge for the purpose of comparing and contrasting the accuracy and acceptability of multiple social-emotional and behavior screeners.

**Study Procedures:** We are requesting your assistance in this study on universal screeners. We will provide a brief training on the study procedures and all materials before the study begins. The study will involve three data collection periods. You will be asked to complete several brief screeners on all students in your class approximately 6 weeks into the school year. On these screeners, you will rate each student's social-emotional skills, behavior and academic performance at school. One month later, you will complete one of the screeners again and fill out a short questionnaire indicating your opinion of the screeners. At the end of the semester we will collect information on students' office discipline referrals (ODRs), suspensions, attendance, and academic grades in core academic subjects. To link student information across data collection periods, you will be sent a spreadsheet with anonymous codes to be assigned to each of your students, which will be used on the screeners in lieu of student names. Once IDs are assigned, the spreadsheets will be password protected by you and resent to the researcher to manage. The research will not have access to your passwords. Details on this procedure will be provided at the training prior to the first data collection. Completing the training, rating scales, and additional student information will take approximately 6-7 hours of your time over the course of the semester. This study is being conducted with your administrator's approval. Frank Gresham, Ph.D. and Shelby Byrd, M.S. 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 social-emotional and 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 schools. In addition, to show our appreciation for your time, effort, and assistance in our research efforts, we will provide each participating teacher with a \$10 gift card. All participating teachers will also be entered into a raffle to win a \$100 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 or fatigued from completing ratings of students' behavior. However, you will be trained on rating procedures to minimize these risks. Furthermore, data will be kept completely confidential through the use of ID numbers, so that data cannot be linked to names.

**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 participation at any time without penalty.

**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 Shelby Byrd at sbyrd5@lsu.edu, Monday-Friday 8:00 a.m. – 4:30 p.m. If you have any questions about participants' 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: \_\_\_\_\_

## APPENDIX I IRB APPROVAL



### ACTION ON EXEMPTION APPROVAL REQUEST

**TO:** Shelby Byrd  
Psychology

**FROM:** Dennis Landin  
Chair, Institutional Review Board

**DATE:** July 13, 2017

**RE:** IRB# E10533

**TITLE:** An Examination of the Technical Adequacy, Classification Accuracy, and Usability of the SSIS SEL Edition Screening/Progress Monitoring Scales in Elementary School Populations

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](mailto:irb@lsu.edu)  
[lsu.edu/research](http://lsu.edu/research)

**New Protocol/Modification/Continuation:** New Protocol

**Review Date:** 7/13/2017

**Approved**  **Disapproved**

**Approval Date:** 7/13/2017 **Approval Expiration Date:** 7/12/2020

**Exemption Category/Paragraph:** 1; 2b

**Signed Consent Waived?:** No

**Re-review frequency:** (three years unless otherwise stated)

**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. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

\* 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>



## VITA

Shelby McCoy Byrd, a native of Milledgeville, Georgia, received her Bachelor of Sciences degree in Psychology from Georgia College & State University in Milledgeville, Georgia. Following graduation, she entered graduate school in the Department of Psychology at Valdosta State University in Valdosta, Georgia. She received her Master of Science degree from VSU in June 2014. Shelby entered the LSU School Psychology doctoral program in August 2014. She is currently completing her predoctoral internship with the Nebraska Internship Consortium in Professional Psychology at Boys Town in Omaha, Nebraska. She expects to graduate with her doctorate in August 2019, and will be completing her postdoctoral fellowship at Boys Town starting in August 2019.