The University of Southern Mississippi

The Aquila Digital Community

Fall 12-2022

Stability of Universal Screening Over Time: An Examination of the Student Risk Screening Scale

Rebecca Wagner Lovelace

Follow this and additional works at: https://aquila.usm.edu/masters_theses

Part of the Elementary Education Commons, and the School Psychology Commons

STABILITY OF UNIVERSAL SCREENING OVER TIME: AN EXAMINATION OF

THE STUDENT RISK SCREENING SCALE

by

Rebecca Wagner Lovelace

A Thesis
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts

Approved by:
Dr. Crystal N. Taylor, Committee Chair
Dr. Zachary LaBrot
Dr. R. Alex Smith

December 2022

Published by the Graduate School

THE UNIVERSITY OF
SOUTHERN
MISSISSIPPI®

ABSTRACT

Universal screening is a proactive method to identify students that are at risk for social-emotional and behavior (SEB) problems and provide information to schools to support early intervention for at risk children. Current recommendations for practice indicate screening should be conducted at three time points during the school year. Previous studies suggest that this recommendation is not empirically based and fewer screenings per year may be sufficient for identifying students at risk. The current study seeks to extend the literature regarding the stability and consistency of screening scores over time by analyzing ratings from the *Student Risk Screening Scale* (SRSS) collected during fall and spring within an urban, elementary school to determine the number of screenings necessary to identify at-risk students. Additionally, this study examined the impact of re-screening students that had borderline risk scores in the fall to determine if a more targeted screening might be feasible in the spring when using the SRSS. Results will inform best practice recommendations related to the implementation of universal screening for SEB concerns using the SRSS in elementary students.

Keywords: Universal Screening, Social-Emotional, Behavior

ACKNOWLEDGMENTS

DEDICATION

This thesis project is dedicated to my family, because this project would not have been possible without their support. I would like to extend a special thanks to my husband, Aaron Lovelace, and my parents Duane and Charlotte Wagner, for their continual prayers and unwavering encouragement and support as I pursue my goals. You may not always know what I'm talking about, but you are always there to listen, strengthen my confidence, and remind me to keep going. I am also sincerely grateful to Michael, Kimberly, Kenny, Renee, Allie, Emma, Papaw, and to my late Nana who have never doubted me; but encourage me, pray for me, and often remind me that taking time to rest is just as important as hard work. Finally, I am ultimately grateful to God for His peace, strength, and for working things out in His own timing along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I - INTRODUCTION

Stability of Universal Screening Over Time: An Examination of the Student Risk

Screening Scale

Social emotional and behavioral (SEB) risk, as defined in previous literature, is

any behavior or behavioral difficulty that impacts the functioning or quality of life for an

individual or their direct caretakers (e.g., teachers, parents, or peers). This definition

includes internalizing behaviors such as social withdrawal and anxiety; externalizing

behaviors such as aggression and disruptive behaviors; inattentive behaviors; and social

skills deficits. SEB risk can vary widely in intensity and severity (Miller et al., 2019). In

in the past 20 years, the prevalence rate of mental health problems in children has

increased substantially (Perou et al., 2013). More recently, the American Academy of

Pediatrics (AAP), the American Academy of Child and Adolescent Psychiatry (AACAP)

and the Children's Hospital Association (CHA) declared a mental health crisis for

children and adolescents in the United States, citing stress from the COVID-19 pandemic

and racial injustice across the country leading to a higher prevalence of mental health

problems and an increased need for service delivery (AAP, 2021). Within this

declaration, there is a call for increased funding for school-based mental health services,

though a push for school-based service delivery has been ongoing since 2009 (AAP,

2021; Splett et al., 2018).

Access to mental health services is not always easy for some individuals. School-

based service delivery models can reduce common barriers associated with accessing

mental health services, such as transportation to sessions and session cost. Multi-Tiered

Systems of Support (MTSS) such as Positive Behavior Intervention and Supports (PBIS)

are prevention models designed to identify students who need SEB support (Sanchez et al., 2018; Splett et al., 2018). PBIS is a three-tiered system in which students receive increasing levels of SEB support based on their needs. Within the model, all students at the Tier I level receive the same level of behavioral supports in the classroom. Those students that do not respond (i.e., need additional support) receive further, targeted service delivery at the Tier II level. Finally, those at the Tier III level that are not responding to targeted, Tier II supports receive more intensive behavioral intervention (Center on PBIS, 2022). An important feature of PBIS is that this model encourages schools to provide interventions to students that are identified as not responding at the Tier I level, which greatly reduces the number of students requiring highly involved, resource intensive supports by intervening early (Bruhn et al., 2013).

Identification Methods

To better reach students that are not responding at Tier I, schools need to be able to identify those students that are not responding. Currently, legislature does not regulate *how* schools identify at risk students, only that schools provide supports to students that need them. Therefore, schools often identify students by using teacher referrals and office discipline referrals (ODRs; Eklund & Dowdy, 2014; Miller et al., 2015). The use of teacher referrals and ODRs as a way for schools to identify at-risk students requires a student to have SEB problems for an extended period of time or have behaviors that are so severe that they warrant immediate support, which is often referred to as a reactive identification method (Glovers & Albers, 2007; Eklund & Dowdy, 2014; Miller et al., 2015). Universal screening, on the other hand, is a more proactive identification method,

but requires more intensive, district-level supports to implement and therefore is used less often (Glover & Albers, 2007; Briesch, et al., 2021).

Eklund and Dowdy (2014) compared the use of a universal screener (*Behavioral Emotional Screening System*; BESS) to identify students at risk for SEB problems to traditional school referral methods and found that of 156 students screened, approximately two thirds identified as being at risk by the screener had not been identified by traditional school referral methods. Similarly, Eklund et al. (2009) compared universal screening to teacher referral and found that over half of the students identified as at risk by the universal screening measure were not identified by teacher report, suggesting that referral systems are not strong methods for finding at-risk students. Teacher referrals are doubly problematic as they are not only reactive in nature, but are susceptible to bias, especially toward minority students and students of low socioeconomic status (Raines et al., 2012).

Similar to teacher referrals, ODRs are reactive in nature. ODRs are typically given once a student's behavior has already escalated to the point of receiving discipline from an administrator in the building. Research does not support using ODRs as the only method of identification (Eklund & Dowdy, 2014; Miller et al., 2015). For example, studies have found that while ODRs may be helpful for providing details about problem behaviors, such as how often it occurs or patterns of behavior, this information is not sufficient on its own to identify students in need of support accurately and before problems begin to escalate (Bruhn et al., 2013; Naser et al. 2018).

Universal Screening for SEB Risk

Despite the research regarding teacher referrals and ODRs as reactive and inefficient tools for identifying at-risk students, they are a commonly employed method of identification within schools. Unlike the 'wait to fail' models of identification, universal screening is a proactive method of identifying students who are at risk of experiencing behavioral difficulties and those data provide schools information necessary to deliver intervention supports to at risk students (Glovers & Albers, 2007).

Being proactive is a clear advantage of universal screening, but this advantage is often outweighed by burdens and barriers that need to be addressed within the universal screening literature. Recent rates of universal screening in schools range from 9-12% for SEB screening as compared to 81% for academic screening and 70% for physical health concerns and this is likely attributed to the lack of explicit, evidence-based guidance (Briesch, et al., 2021; Bruhn et al., 2014; Lane et al., 2015). School personnel report that referral methods of identification require significantly less resources and time compared to gathering universal screening data (Briesch et al., 2021; Bruhn et al., 2014; Naser et al., 2018). Additionally, some studies have suggested that the underuse of SEB screening can be attributed to administrators and other stakeholders lacking the knowledge about the importance or even existence of screening for SEB risk (Volpe & Briesch, 2018; Briesch et al., 2018). Likewise, many school districts cite a lack of guidance regarding the implementation of a screening protocol in schools (Burns & Rapee, 2021).

The lack of consistent and explicit guidance for screening procedures inhibits school districts from exploring universal screening as an option (Volpe & Briesch, 2018; Briesch et al., 2018). Briesch et al. (2018) conducted a review of state-level guidance

4

regarding universal screening that was produced by each state's department of education. They found that approximately half of the states did not provide explicit directions regarding screening practices, with nine states not mentioning SEB screening at all. Other states mentioned SEB screening only in the context of MTSS, including recommending the use of ODRs as a screener or only screening when specific issues arise, such as academic difficulties or behavioral issues such as truancy prevention, substance use, and eating disorders. For states that did explicitly discuss SEB screening, recommendations varied regarding the frequency of screening. Variations ranged from screening 2-4 times per year to only screening certain grades or only screening as part of school readiness determinations. At a federal level, legislature such as Every Student Succeeds Act (ESSA, 2015) and Individuals with Disabilities Education Act (IDEA, 2004) include verbiage that recognizes the importance of early identification and focuses the on 'whole child', but do not provide the guidance necessary for schools to make decisions regarding which screener to use, how to use the screener, or when to use the screener.

Current Guidance for Universal Screening Practices

To date, most of the advances in the literature have focused on establishing reliable and valid measures for screening. There have been several measures identified in the literature with psychometric support, such as the *Student Risk Screening Scale* (SRSS; Lane & Menzies, 2009), the *Social, Emotional, and Behavior Risk Screener* (SAEBRS; Kilgus & von der Embse, 2014)*,* and the *Behavioral Emotional Screening System* (BESS; Kamphaus & Reynolds, 2015)*.* While establishing psychometrically sound measures is important groundwork for universal screening practices, it is not enough to provide school personnel with the information they need to implement screening effectively.

Further research to support best practices related to the screening process are vital to establishing evidence-based screening procedures and addressing the research to practice gap that exists for universal screening for SEB risk (Volpe & Briesch, 2018).

It is important for researchers to establish evidence-based practices so that school districts can implement universal screening with little to no support from outside agencies just as they are with referral systems from teachers and ODRs. Current calls for research regarding universal screening focus on establishing evidence-based and consistent guidelines for practice related to SEB screening. Areas of focus include general procedures for screening; including research regarding how to select measures, informants, and interventions; how to use the date effectively once collected, how to minimize costs related to screening, how to maximize student benefit, and how to determine the frequency of screening (Volpe & Briesch, 2018; Cook et al., 2010).

One area of universal screening research that Volpe and Briesch, (2018) and Cooke et al., (2010) indicated was necessary within the universal screening literature is research identifying the frequency with which schools should be implementing universal screening procedures. Publishers for the SRSS, SAEBRS, and BESS have varying recommendations regarding the frequency of universal screening when using their measures. The recommendations for the SRSS are to screen three times per year with a two-week window to complete the screening (Rollenhagen et al., 2021). The SAEBRS does not provide a fixed number of times to screen but indicates students can be screened up to five times per year (Illuminate Education, n.d.), and the BESS does not provide any recommendations about how often screening should be completed (Pearson, n.d.).

Additionally, there are free resources available online that provide school districts with best practice recommendations for using universal screeners in general (see Romer et al., 2020). These general recommendations for screening frequency state three times per year is best practice, but this schedule has not been empirically evaluated for effectiveness (Walker et al., 2014; Romer et al., 2020) and is likely guided by recommendations made for academic screening (Volpe & Briesch, 2018; Dowdy et al., 2014). However, even within academic screening literature, screening three times a year may not be necessary. VanDerHeyden et al. (2018) examined the stability of universal screening scores over time and the efficacy of targeted screening of reading proficiency and found that most students who were at risk at the first screening remained at risk at subsequent time points, suggesting that a single administration of screening measures may be sufficient to identify students in need of reading interventions. Thus, further indication that recommendations regarding the frequency of universal screening needs to be empirically studied.

Research Regarding the Frequency of Universal Screening

Researchers have begun to examine the stability of SEB risk over time to inform recommendations for the frequency of screening administration. To date, there are four studies that have specifically looked at the stability of universal screening scores for SEB risk over time.

Dowdy et al. (2014) conducted a study evaluating the long-term stability of screening scores over time for 156 students. Data were collected across four years while participants were in 8th to 11th grade. In this study, researchers used the self-report version of the BESS and screened students three times annually (fall, winter, and spring)

to examine stability of scores over time. They found that overall, risk was stable across time points, with the majority of students (about 90%) initially identified as "not at risk" remaining "not at risk" at subsequent screenings. Additionally, they found that students in the elevated range tended to decrease in risk over time, with about two-thirds of students initially identified as having elevated risk falling into the normal range at the final screening period. The results of this study supports the stability of social and emotional behaviors over time because the majority of students remained the same or decreased in risk over time, which may support less frequent screening periods.

Kilpatrick et al. (2018) conducted a study evaluating the Devereux Student Strengths Assessment (DESSA)-mini with the primary purpose of investigating the frequency of screening for SEB risk and stability of SEB risk scores over time. This study examined screening data for early elementary students (K-2) and contained data for 525 participants across all three grades. The results of this study provided support for annual screening, with fall screening scores being equally correlated to outcome variables compared to winter and spring screening scores. Risk status was especially stable from winter to spring screening sessions, suggesting annual or bi-annual screening may be sufficient. They also found risk status was stable over time, with most changes in risk status occurring in students moving from at-risk to not at risk. One large limitation this study found was that the DESSA-mini was not a reliable measure and demonstrated poor diagnostic accuracy of SEB risk.

Dever et al. (2018) conducted a longitudinal study evaluating the BESS with an early childhood population to evaluate the stability of risk over time and determine if a targeted screening procedure (e.g., re-screening only a portion of students using a specific

criterion, such as a cutoff score) would reduce the need for multiple screenings. They evaluated ratings from 230 teachers for 1,014 preschool, kindergarten, and first-grade students. The results of their study showed that risk status was stable from fall to spring screenings, but was not stable from year to year, suggesting that screening at least once per year is necessary to identify students in need of SEB supports. Additionally, they were able to determine a cut score that could be used with future screening administrations to create a gating procedure for re-screening. Since this study was the first to evaluate a cut score for use with a gating procedure, further research is needed to determine the generalizability to broader student populations and to determine the use of this strategy with other screening measures.

The most recent study, conducted by Miller et al. (2019), aimed to evaluate the stability of SEB risk screening scores over time on three different screening measures. They also investigated the efficiency of only re-screening students whose risk scores fell in what was considered a 'borderline' range as opposed to rescreening all students, which would significantly reduce the volume of resource allocation needed for additional screening sessions. To evaluate stability, students were rated at three time points on three different screening measures, the Direct Behavior Rating – Single Item Scales (DBR – SIS), the BESS, and the Social Skills Improvement System – Performance Screening Guide (SSIS – PSG). Ten students per participating classroom were selected for analysis. Data from 1,964 students across 1st, 2nd, 4th, 5th, 7th, and 8th grades were used. Findings indicated that the BESS was the most conservative measure and produced the most stable risk classification compared to the other measures. Risk over time had similar results to the Kilpatrick et al. (2018) study, with risk being relatively stable over time and the most

stable period occurring from winter to spring. This result further supports that screening

twice per year may be adequate to identify students whose risk status changes over the

school year. For students who fell in the borderline range, re-screening in winter

following the initial fall screening produced the greatest results. Overall, targeted

screening for students in borderline ranges produced varied results that require follow-up

to determine the feasibility of targeted screening. Though this information is helpful, this

study is not without limitations. In particular, the researchers compared three different

measures that used three very different forms of determining risk (e.g., norm-referenced

vs. criterion referenced vs. observation ratings) and all teachers and students were not

included in the study (e.g., only certain grades and only 10 students per classroom were

used).

Evaluating the most effective schedule for screening is important information for

schools to be able to allocate resources for screening. If fewer screenings per year

produce similar results to screening three times per year, it could improve the screening

process as a whole and potentially remove some barriers to screening for schools, such as

cost and resource allocation, including time dedicated to analyzing screening data and

intervention selection and implementation (Kilpatrick et al., 2018). Current research has

examined the frequency of universal screening using the BESS, DESSA-mini, DBR-SIS,

and the SSIS and have determined that fewer screenings might be sufficient for

identifying student that need additional Tier II and Tier III supports. However, another

commonly used universal screening measure, the SRSS has not been empirically studied.

Purpose

The current study aims to extend the current literature by evaluating the stability of screening scores on the SRSS across time points. The SRSS is a psychometrically sound universal screening instrument that provides scores that fit within PBIS models. Ratings from the SRSS yield two scores, one for externalizing behaviors and one for internalizing behaviors (Lane et al., 2015). Scores are then coded as low risk, moderate risk, or high risk, which corresponds to the three tiers of a PBIS system. Given the likelihood that schools will choose to use a measure that fits within their current MTSS system, there is a need to examine the frequency of screening for this measure in particular.

Fall and spring universal screening data, using the SRSS, from an urban elementary school population were used to identify the frequency with which screening should occur for schools that use the SRSS. The internalizing and externalizing domains were assessed separately. It was hypothesized that the results of this study will provide further evidence for best practices related to screening frequency for SEB risk and stability of risk status across time points when using the SRSS. The following research questions are of interest in the current study.

1. How stable is internalizing risk status for students across time points?
2. How stable is externalizing risk status for students across time points?
3. What proportion of students identified as at risk for internalizing behaviors in the spring were not identified in the fall?
4. What proportion of students identified as at risk for externalizing behaviors in the spring were not identified in the fall?

5. Does identifying students in the borderline range for internalizing risk in the fall

    reduce the need for spring screening?

6. Does identifying students in the borderline range for externalizing risk in the fall

    reduce the need for spring screening?

CHAPTER II – METHODS

Participants and Setting

A de-identified data set of SRSS ratings collected as part of regular school
practice was used for analysis. District-level and IRB approval was secured before data
were analyzed. Screening scores were collected during the 2021-2022 school year in an
urban school district in the southeastern United States. Based on report from the State
Department of Education, 87 percent of students in the district are Black or African
American and seven percent are Hispanic or Latino. The school district is considered
Title 1 and all students receive free or reduced-price lunch. The data set contained scores
from five elementary schools (grades K-5) with ratings from 103 teachers for 1,896
students. See Table 1 for a breakdown of students and teachers by grade. Demographic
data were not collected on participants because this analysis occurred following the
screening window and demographic data were not collected as part of typical practice.

Measures

The Student Risk Screening Scale

The *Student Risk Screening Scale* (SRSS, Lane & Menzies, 2009) is a teacher-
report universal screening tool. Ratings are split into two domains, internalizing and
externalizing behaviors, and scores produce a risk level (low, moderate, or high) for each
domain. Teachers rate all behaviors on a Likert scale ranging from 0-3 where 0 = *never*
and 3 = *frequently*. The items on the internalizing scale are emotionally flat; shy,
withdrawn; sad, depressed; anxious; and lonely. For internalizing behaviors, ratings
between 0-1 are considered low risk, scores of 2 or 3 are considered moderate risk, and
scores ranging from 4-15 are considered high risk. The items on the externalizing scale

are steal; lie, cheat, sneak; behavior problem; peer rejection; low academic achievement; negative attitude; and aggressive behavior. Ratings between 0-3 are considered low risk, scores ranging from 4-8 are considered moderate risk, and scores between 9-21 are considered high risk for externalizing behaviors. Previous research supports the reliability of the SRSS with elementary populations with strong internal consistency (>0.80) and correct classification rates of 0.81 (Lane et al., 2015).

<div align="center">Procedures</div>

Universal screening occurred as part of typical practice within the school district. Screening was facilitated by two doctoral level graduate students in school psychology and overseen by a supervising professor in school psychology and the director of the special education department within the school district. Screening was conducted in the fall and spring during the 2021-2022 school year, with the first administration occurring four weeks into the fall semester and the second administration occurring four weeks into the spring semester. The same method was employed for each administration. School counselors were responsible for distributing screening instructions to the elementary school teachers. All counselors attended an administrative meeting where the school psychology graduate students gave a presentation about the importance of universal screening and explained each step of the procedure for completing the screening sheets. Presentations also included a definition of externalizing and internalizing behavior and emphasized the importance of completing the measure accurately. Immediately following the presentation to the counselors, the graduate students shared a Google Drive folder containing the screening sheet for each school with the appropriate counselor. All sheets were labeled by teacher name, and counselors were asked to share the appropriate sheet

with each teacher and provide the presentation PowerPoint that included definitions of externalizing and internalizing behavior, instructions for completing the measure, and a slide emphasizing the importance of completing the measure on time and accurately. Teachers were given two weeks to complete the screening measures. At the one-week mark, the graduate students emailed all counselors and principal to remind them of the screening deadline. The email contained the instructions for completing the screening measure and there was a note to contact the graduate students if the counselor or teacher had any questions or needed help completing the screener.

Following the spring screening in 2022, researchers reached out to the school district to request the de-identified SRSS data from the 2021-2022 school year. SRSS data from fall 2021 and spring 2022 were provided to the researchers following approval from the university's Institutional Review Board. Data were provided to researchers through a password protected Google Sheets document for each school with a tab for each teacher. School names, teacher names, and student names were removed. Within each teacher's tab, an excel sheet displayed each student's grade level, SRSS internalizing score, SRSS internalizing category, SRSS externalizing score, and SRSS externalizing category. See Appendix C for an example of the Google Sheet document researchers received from the school district. Researchers created random student ID numbers that connected each student with their classroom and school. These data were then copied and pasted into an SPSS document for statistical analysis.

Data Analysis

Prior to data analysis, the researcher examined the data to look for missing cases (e.g., any behavior ratings that were left blank, or a blank total score for either domain),

rating errors (e.g., providing ratings outside of the 0-3 range), and students who did not have complete data (e.g., ratings for both domains at both time points). Individuals with missing cases, rating errors or incomplete data were excluded from the analysis (see Miller et al., 2019). The original data set contained ratings from 103 teachers for 1,896 students across all five elementary schools. Six teachers were excluded for being out of grade range (e.g., they taught Pre-K and the elementary version of the SRSS is not validated for use with a Pre-K population). To analyze the stability of scores over time, only students with complete screening data (ratings at both time points) were included in analyses. This method was based on previous studies that have evaluated stability of screening scores over time and used complete cases rather than imputation since imputation can be biased and is not appropriate for assessing stability and patterns of risk (Miller et al., 2019; Dever et al., 2018). Three teachers were excluded for not having any screening data (e.g., the teacher did not complete the screener for their class or completed the entire screener incorrectly, such as providing ratings outside the 0-3 range or not rating the behaviors at all). An additional 12 teachers were excluded for only having fall or spring data, but not both. The final number of teachers included in the analysis was 82 across all schools and grades (see Table 1 for a breakdown of teachers and students by grade). A total of 1,637 students remained in the sample across all five elementary schools. Individual students were then excluded for only having fall or spring data, but not both, resulting in 415 students being removed from the sample. An additional 53 students were excluded for having incomplete or incorrect data (e.g., the teacher did not rate all items on the screener or used out of range ratings) at either time point. The final number of students included in the analysis was 1,169. Following exclusions, a numerical

value was assigned to risk variables to replace categorical labels, where 1 = low, 2 = moderate, and 3 = high risk for both externalizing and internalizing risk. These numerical values were designated by the researcher for the sole purpose of this study to simplify SPSS analyses. For internalizing behaviors, ratings between 0-1 are considered low risk, 2-3 is considered moderate risk, and 4-15 is considered high risk. While ratings between 0-3 are considered low risk, 4-8 is considered moderate risk, and 9-21 is considered high risk for externalizing behaviors.

Once the data were prepared, group statistic analytic approaches were used to evaluate the stability of ratings across time points and assess identification of students with scores in the borderline range. First, descriptive statistics were generated to assess the number of students at risk in each category at each time point. To analyze stable risk patterns across time points, phi correlations (or stability coefficients) were generated for each risk category. To analyze changes in risk over time, Cohen's kappa was used to analyze consistency of scores for internalizing and externalizing risk, with each time point (e.g., fall and spring) treated as a rater. Cohen's kappa ranges from -1 to 1, where 1 indicates perfect agreement between raters. Values equal or less than 0 indicate no agreement, 0.01-0.20 indicates none to slight agreement, 0.21-0.40 indicates fair agreement, 0.41-0.60 indicates moderate agreement, 0.61-0.80 indicates substantial agreement, and 0.81-1.00 indicates almost perfect agreement (McHugh, 2012).

The final analysis examined borderline risk status to identify students that may need to be re-screened in the Spring. Borderline scores were determined to be those total scores on the SRSS that ranged between two and three for externalizing behavior and scores of one were determined to be borderline for internalizing behavior. Borderline

scores for externalizing and internalizing risk were selected because scores between the

low and moderate range and the moderate and high were small (e.g., one point in some

instances). For example, ratings for externalizing behavior have cut scores of 0-3 for low

risk, 4-8 for moderate risk, and 9-21 for high risk. So, borderline risk for externalizing

behavior was defined as a score between 2-3. Similarly, for internalizing behavior, scores

of 0-1 are considered low risk, 2-3 is considered moderate risk, and 4-15 is considered

high risk. So, borderline risk for internalizing behavior was defined as a score of 1. All

data were analyzed using SPSS version 22.

CHAPTER III - RESULTS

Descriptive statistics for the percentage of students at risk in each category at each time point can be found in Table 2. During the fall screening, 15.7% of students screened were identified as being moderately at-risk for externalizing concerns, and 10.9% of students screened were identified as being moderately at-risk for internalizing concerns. Regarding high risk, 5.6% of students screened were identified as being high risk for externalizing concerns, and 6.7% were identified as being high risk for internalizing concerns. During the spring screening, 17% of students were identified as moderate risk for externalizing concerns and 10.9% were identified as being moderately at risk for internalizing concerns. Further, in the spring, 5.6% were identified as being high risk for externalizing concerns, and 6% were identified as being high risk for internalizing concerns. *Figures 1 and 2* show the percentages of each student in each risk category and for those in the borderline range at each time point.

Research Questions 1 and 2: Stability of Risk Status for Students Across Time Points

Stability coefficients across time ranged from ($\Phi = 0.46$ to 0.59), indicating a moderate, positive relationship of risk status between fall to spring for externalizing and internalizing concerns. Specifically, stability for those found to be at high risk or moderate risk for externalizing behavior remained strong between fall and spring ($\Phi = 0.59$). Regarding those at moderate or high risk for internalizing behavior, coefficients indicated a weaker relationship between fall and spring ($\Phi = 0.46$). Phi correlations for risk status across time points can be found in Table 3.

Questions 3 and 4: Proportion of Students Identified as At-Risk in the Spring, Not Fall

Cohen's kappa was used to determine consistency across scores for externalizing and internalizing behavior from fall to spring. Results showed low consistency across time points (externalizing κ = 0.39, internalizing κ = 0.33). This can also be interpreted as only 39% and 33% of students who were screened in the fall remained at the same level of risk for externalizing and internalizing concerns, respectfully, in the spring.

Furthermore, crosstabulations for externalizing risk status showed that of those students identified as low risk in the fall, 87.6% remained low risk at spring screening while 11% were identified as moderate risk, and 1.4% were identified as high risk at spring screening. For those identified as moderate risk for externalizing problems, 40.4% of those identified in the fall remained moderate risk at spring screening while 47.5% were identified as low risk in the spring, and 12% were identified as high risk in the spring. For externalizing high risk, 47% remained high risk at spring screening, 36.4% were identified as moderate risk at spring screening, and 16.7% were identified as low risk at spring screening.

Crosstabulations for internalizing risk status showed that of those identified as low risk in the fall, 90.1% remained low risk at spring screening while 6.9% were identified as moderate risk, and 3% were identified as high risk at spring screening. For internalizing moderate risk, 29.7% of those identified in the fall remained moderate risk at spring screening while 60.2% were identified as low risk in the spring, and 10.2% were identified as high risk at spring screening. For those identified as having high internalizing risk, 35.9% of those identified in the fall remained high at spring screening

20

while 29.5% were identified at moderate risk, and 34.6% were identified as low risk at spring screening.

<p align="center">Questions 5 and 6: Impact of Borderline Screening in the Spring</p>

To analyze the possibility of re-screening *only* those students that were in the borderline range in the fall, the stability and consistency of ratings for those in the borderline range was examined. The borderline range was determined to be scores on the SRSS that fell in the 2-3 range for externalizing behavior and scores of 1 for internalizing behavior. Scores were coded in SPSS as 0 for no risk, 1 for borderline risk, and 2 for at-risk. At risk included those in the moderate and high ranges on the SRSS. The total percentage of students in the borderline category for externalizing behavior in the fall was 22.5% and the total percentage of students in the borderline category for internalizing behavior in the fall was 10.4%. Of those identified as borderline risk for externalizing behaviors in the fall, 39.9% remained borderline at spring screening while 25.9% additional students were identified as at-risk, and 34.2% were identified as not at risk at spring screening. Of those identified as borderline risk for internalizing behaviors in the fall, 18% remained borderline at spring screening while 18.9% were identified as at-risk, and 63.1% were identified as not at risk at spring screening.

CHAPTER IV – DISCUSSION

Recent calls for research regarding universal screening for SEB risk have focused on establishing guidelines for screening practice, including the frequency of screening administration (Volpe & Briesch 2018; Cook et al., 2010). Previous studies have examined the stability of universal screening scores over time using measures such as the SSIS, the BESS, and the DBR-SIS, but to date none have examined the SRSS. Across these studies, results showed that SEB risk was relatively stable over time, but not to the extent where a single screening administration per year was supported (Dever et al., 2018; Dowdy et al., 2014; Miller et al., 2019). Additionally, previous literature has examined the use of a targeted screening procedure where only students falling in a borderline risk range are rescreened to reduce the resources needed for screening. The purpose of this study was to address the gap that exists in the current literature by evaluating the stability and consistency of risk status across time points with the SRSS, as this measure has not been evaluated in previous studies. Additionally, this study examined the use of a targeted screening approach where a second screening in the spring might be used for only those identified as at-risk within a specified borderline range.

Stability of Risk Over Time

Phi correlations were used to evaluate the stability of risk over time and showed a moderate level of consistency of ratings from fall to spring. The highest consistency occurred for borderline externalizing risk status. The lowest amount of consistency occurred for internalizing behavior, which can be attributed to a significant number of students who were identified as having moderate risk in the fall being categorized as low risk in the spring. This sharp decrease in at-risk students could be explained by teachers

22

having more time to become acquainted with students and increased ability to accurately detect the selected behaviors. Fall screening typically occurs 4-6 weeks into the semester, so teachers may not be fully acquainted with students or may be struggling to adjust to a new group of children - inflating scores. By spring, teachers have had time with their students and may be able to more accurately rate behavior (barring students remain in the same classroom all year). It is also possible that behavior supports available to the student changed their observable behavior, impacting teachers rating of risk. Data regarding PBIS implementation and tiered behavioral supports was not available for this sample, however, based on report from the school PBIS was not implemented consistently across each elementary. Therefore, these results are like those found in previous literature, which cite moderate consistency as being indicative of spring screening identifying novel risk, providing support for screening at least twice per year in fall and spring (Miller et al., 2019).

Consistency of Scores

For students who were identified as having low risk at the fall screening for externalizing behaviors, most remained at low risk during spring screening. Similarly, most students who were identified as having low risk at the fall screening for internalizing behaviors remained at low risk during spring screening. For both externalizing and internalizing risk, most of the movement occurred from low to moderate risk. For moderate risk status, many students identified during fall screening were rated as being low risk at spring screening. Consistency was low between fall and spring, indicating there is a need for at least one more screening period outside of the one in the fall to properly identify students in need of support.

Similarly, these differences in risk status could be due to teachers having more time to get acquainted with the student or seeing improvement in students who are receiving supports during the school year. It also could be possible that when spring screening occurs, teachers are burnt out or overwhelmed which may lead to less integrity completing the measure, including filling the measure out with all zeros leading to more students identified as low risk (Miller et al., 2019; Dever et al., 2018). Future research in this area is needed to determine the impact of teacher perception on risk status.

<div align="center">Borderline Screening</div>

To explore the feasibility of a targeted screening approach, where only those students that fall within a specified range are re-screened in the spring, students were organized into three categories, at-risk (including moderate and high risk), borderline risk, and not at-risk categories. Statistics were run to see if re-screening students in the borderline risk group would be sufficient for identifying novel risk. Students in the borderline range were chosen because it is hypothesized that students falling in borderline risk categories may be more likely to move up to 'at-risk' than those who are not at-risk, based on Miller et al., (2019).

For borderline categories, approximately 33% of students fell in the borderline range for either internalizing or externalizing concerns during the fall screening, which means only re-screening this group of students in the spring would require significantly less resources than rescreening all students. In this sample, nearly 26% of the students that were in the borderline range for externalizing behavior risk in the fall were at risk in the spring while around 40% remained in the borderline range in the spring. For internalizing behaviors, 19% of those in the borderline range in the fall were newly

<div align="center">24</div>

identified as at-risk during spring screening and nearly the same percent of students

(18%) remained in the borderline range in the spring. Overall, a large proportion of

students remained in the borderline range for externalizing concerns but only a small

proportion remained in the borderline range for internalizing concerns. Additionally,

between 19% and 26% of students that were in the borderline range would have been

identified as at risk if screening only occurred in the fall. Additional research to

investigate targeted screening for borderline risk status is needed to evaluate the

appropriateness of this approach, but it might be feasible to only re-screen those in the

borderline range for externalizing behavior concerns.

<center>Limitations and Future Research</center>

There are several limitations to be discussed related to the current study. First, this

study analyzed data that were previously collected and de-identified by a school district.

Therefore, there is limited information available about the population of teachers and

students that participated in the screening process. In the future, collecting information

about years of teaching, experience and training with universal screening, and general

teacher demographics could be important to evaluate the extent of generalization of the

findings. Related to generalization, this study only examined screening scores on one

measure, the SRSS. Therefore, it is difficult to generalize the results to other screening

measures. Additionally, there is no data available about the current state of PBIS

implementation within each school and the quality of supports provided at each tier.

Future studies could collect information about current behavior supports that may provide

explanation for decreasing risk status across timepoints.

<center>25</center>

Second, almost one-fourth of students were removed from the original sample due to teacher errors that were directly noticeable (e.g., out of range ratings or missing data). It is possible that other errors exist that were not out of range or directly noticeable. For example, there may be errors in completion due to teachers not fully understanding the process and purpose of universal screening, or not understanding what the behavior may look like (e.g., what does a 2 on aggressive behaviors truly mean?). Using complete cases also did not address whether errors were random or on purpose, and future studies may conduct statistical analyses to determine the best way to handle missing and incorrect data.

Third, this study used data that were collected as part of typical practice in a school district. As part of this districts screening procedures, teachers screen students in the fall and spring each year. Other researchers have looked at screening across different timepoints including fall, winter, and spring to determine the number of screenings and the time of year that screening should occur. Future research examining the SRSS may examine the use of the SRSS across all three timepoints to determine when and how often screening should occur.

Finally, the data in this study were collected following a full return to school from COVID-19. While we know there has been a youth mental health crisis declared in light of the pandemic, little else is known about the direct effects of COVID-19 on SEB risk and the impact it might have on universal screening (NASP, 2020). Further research is needed to evaluate the impact on universal screening results.

Conclusion

        Current calls for research regarding universal screening for SEB risk have focused on establishing guidelines for screening practice, including the frequency of screening administration (Volpe & Briesch 2018; Cook et al., 2010). Few studies have examined the stability of universal screening over time. Furthermore, few studies have evaluated the use of a targeted screening approach where only students in a borderline range are re-screened at subsequent administrations to reduce the number of resources needed to implement universal screening in schools and to reduce the likelihood of over screening. This study attempted to fill that gap by examining the stability and consistency of universal screening using the SRSS as well as the use of a targeted screening procedure. Risk status was moderately stable over time, and many students' risk status changed between fall and spring screening administrations. Because of this change in risk status, screening students one time per year may not be sufficient to identify all students with SEB risk, providing support for multiple screening administrations per year. Additionally, novel students were identified when examining the implications for re-screening students that fall within a borderline range and it was concluded that it might be feasible to only re-screen students that are at risk in the borderline range for those with externalizing behavior concerns. Further research is needed to examine the use of borderline screening and its potential use in practice.

APPENDIX A – Tables and Figures

Table A1

Total Number of Teachers and Students Included in the Final Analysis by Grade

Category	Teachers	Students
Kindergarten	14	210
First Grade	13	203
Second Grade	12	199
Third Grade	12	205
Fourth Grade	11	123
Fifth Grade	11	179
Self-Contained	9	50
Total	82	1169

Table A2

Risk Category Descriptive Statistics by Timepoint

Category	Fall	Spring
Externalizing Low	920 (78.7)	904 (77.3)
Externalizing Moderate	183 (15.7)	199 (17)
Externalizing High	66 (5.6)	66 (5.6)
Internalizing Low	963 (82.4)	972 (83.1)
Internalizing Moderate	128 (10.9)	127 (10.9)
Internalizing High	78 (6.7)	70 (6)

Note. Values in parentheses represent percentage.

Table A3

Phi Correlations from Fall to Spring

Category	Phi
Externalizing Behavior	.585
Internalizing Behavior	.464

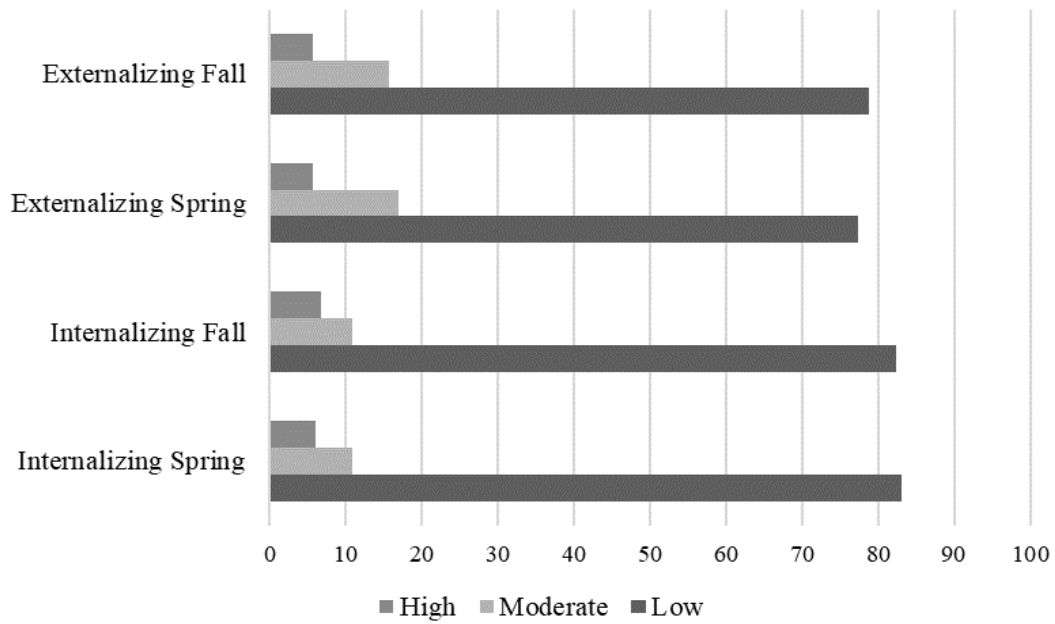*Note*. All correlations were significant at p < .001.

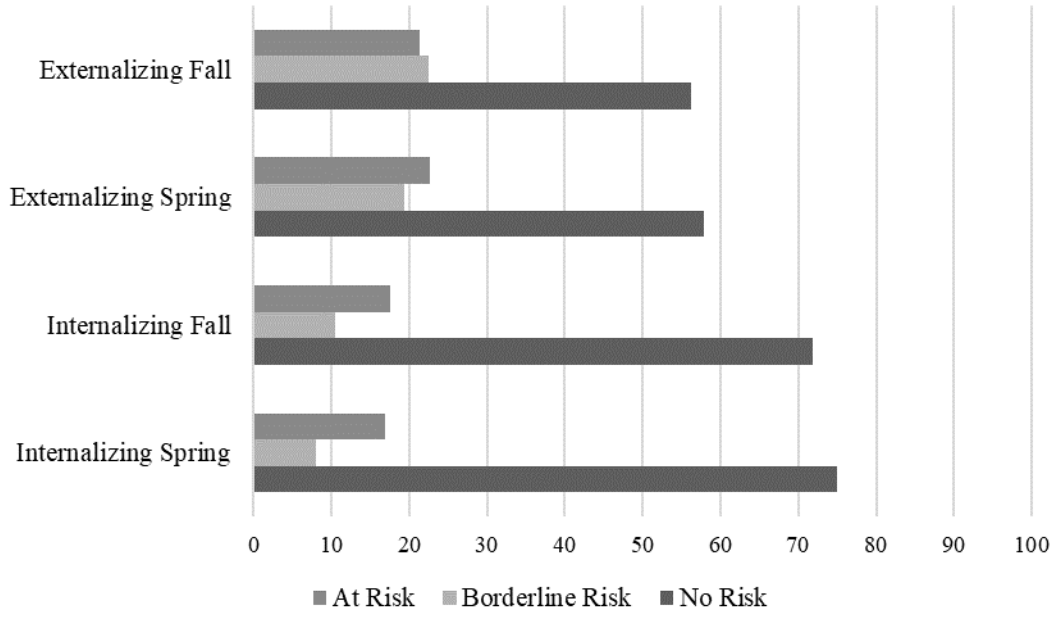Figure A1. *Percentages of students in each SRSS risk category across time*

Figure A2. *Percentages of students categorized as at-risk, borderline risk, or no risk across time.*

APPENDIX B –IRB Approval Letter

Office *of* Research Integrity

THE UNIVERSITY OF SOUTHERN MISSISSIPPI.

118 COLLEGE DRIVE #5116 • HATTIESBURG, MS | 601.266.6756 | WWW.USM.EDU/ORI

NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

The project below has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services regulations (45 CFR Part 46), and University Policy to ensure:

- The risks to subjects are minimized and reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered involving risks to subjects must be reported immediately. Problems should be reported to ORI via the Incident submission on InfoEd IRB.
- The period of approval is twelve months. An application for renewal must be submitted for projects exceeding twelve months.

PROTOCOL NUMBER: 22-852
PROJECT TITLE: Stability of Universal Screening Over Time: An Examination of the SRSS
SCHOOL/PROGRAM   Psychology
RESEARCHERS: PI: Rebecca Lovelace
Investigators: Lovelace, Rebecca~Taylor, Crystal Nichole~LaBrot, Zachary C~
IRB COMMITTEE ACTION: Approved
CATEGORY: Exempt Category
APPROVAL STARTING: 06-Jul-2022

Donald Sacco, Ph.D.
Institutional Review Board Chairperson

31

APPENDIX C Sample SRSS Document

Student ID	Grade	SRSS-E7 Total	SRSS-E7 Risk Category Low Risk, 0-3 Moderate Risk, 4-8 High Risk, 9-21	SRSS-I5 Total	SRSS-I5 Risk Category Low Risk, 0-1 Moderate Risk, 2-3 High Risk, 4-15
	3	13	High Risk	9	Moderate Risk

REFERENCES

American Academy of Pediatrics. (2021, October 19). *AAP-AACAP-CHA declaration of a national emergency in child and adolescent mental health.* https://www.aap.org/en/advocacy/child-and-adolescent-healthy-mental-development/aap-aacap-cha-declaration-of-a-national-emergency-in-child-and-adolescent-mental-health/

Briesch, A. M., Chafouleas, S. M., & Chaffee, R. (2018). Analysis of state-level guidance regarding school-based universal screening for social, emotional, and behavioral risk. *School Mental Health*. Advance online publication. https://doi.org/10.1007/s12310-017-9232-5

Briesch, A.M., Chafouleas, S.M., Dineen, J.N., McCoach, B., & Donaldson, A. (2021). School building administrator reports of screening practices across academic, behavioral, and health domains. *Journal for Positive Behavior Interventions*, *00*(0), 1-12. https://doi.org/10.1177%2F10983007211003335

Bruhn A. L., Lane K. L., Hirsch S. E. (2013). A review of tier 2 interventions conducted within multitiered models of behavioral prevention. *Journal of Emotional and Behavioral Disorders*. *22*(3), 171-189. doi:10.1177/1063426613476092

Bruhn, A. L., Woods-Groves, S., Huddle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education and Treatment of Children*, *37*, 611–634. https://doi.org/10.1353/etc.2014.0039

Burns, J. R., & Rapee, R. M. (2021). Barriers to universal mental health screening in schools: The perspective of school psychologists. *Journal of Applied School*

Psychology. Advance online publication.

https://doi.org/10.1080/15377903.2021.1941470

Center on PBIS (2022). *Positive Behavioral Interventions & Supports* [Website].

www.pbis.org.

Cook, C. R., Volpe, R. J., & Livanis, A. (2010). Constructing a roadmap for future

universal screening research beyond academics. *Assessment for Effective*

Intervention, 35(4)*,* 197-205.

Dever, B. V., Dowdy, E., & DiStefano, C. (2018). Examining the stability, accuracy, and

predictive validity of behavioral–emotional screening scores across time to inform

repeated screening procedures. *School Psychology Review, 47*(4), 360-371. doi:

10.17105/SPR-2017-0092.V47-4

Dineen, J. N., Chafouleas, S. M., Briesch, A. M., McCoach, D. B., Newton, S. D., &

Cintron, D. W. (2022). Exploring social, emotional, and behavioral screening

approaches in U.S. public school districts. *American Educational Research*

Journal, 59(1), 146-179. doi: 10.3102/00028312211000043

Dowdy, E., Nylund-Gibson, K., Felix, E. D., Morovati, D., Carnazzo, K. W., & Dever, B.

V. (2014). Long-term stability of screening for behavioral and emotional risk.

Educational and Psychological Measurement, 74(3), 453-472. doi:

10.1177/0013164413513460

Eklund, K. & Dowdy, E. (2014). Screening for behaioral and emotional risk versus

traditional school identification methods. *School Mental Health, 6*, 40-49.

doi.org/10.1007/s12310-013-9109-1

Eklund, K., Renshaw, T. L., Dowdy, E., Jimerson, S. R., Hart, S. R., Jones, C. N., &
Earhart, J. (2009). Early identification of behavioral and emotional problems in
youth: Universal screening versus teacher-referral identification. *The California
School Psychologist, 14*, 89-95.

Every Student Succeeds Act of 2015, Pub. L. No. 114 –95 § 114 Stat. 1177 (2015).

Glovers, 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.

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 et
seq. (2004).

Illuminate Education. (n.d) *SAEBRS (The Social, Academic, Emotional Behavior Risk
Screener).* https://www.illuminateed.com/products/fastbridge/social-emotional-
behavior-assessment/saebrs/

Kamphaus, R. W. (2012). Screening for behavioral and emotional risk: Constructs and
practicalities. *School Psychology Forum: Research In Practice, 6*(4), 89-97.

Kamphaus R. W., & Reynolds, C. R. (2015). Behavior assessment system for children,
second edition (BASC-2): Behavioral and emotional screening system (BESS).
Bloomington, MN: Pearson.

Kilgus, S. P., & von der Embse, N. P. (2014). Unpublished technical manual of the
Social, Academic, and Emotional Behavior Risk Screener.

Kilpatrick, K. D., Maras, M. A., Brann, K. L., & Kilgus, S. P. (2018). Universal
screening for social, emotional, and behavioral risk in students: DESSA-mini risk

stability over time and its implications for screening procedures. *School Psychology Review, 47*(3), 244-257. doi: 10.17105/SPR-2017-0069.V47-3

Lane, K. L. & Menzies, H. M. (2009). Student Risk Screening Scale for Early Internalizing and Externalizing Behavior (SRSS-IE). Screening scale. Available at Ci3t.org/screening.

Lane, K. L., Oakes, W. P., Swogger, E. D., Schatschneider, C., Menzies, H., M., & Sanchez, J. (2015). Student risk screening scale for internalizing and externalizing behaviors: Preliminary cut scores to support data-informed decision making. *Behavioral Disorders*, *40*, 159-170. https://doi.org/10.17988/0198-7429-40.3.159

McHugh M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276–282.

Miller, F. G., Cohen, D., Chafouleas, S. M., Riley-Tillman, T. C., Welsh, M. E., & Fabiano, G. A. (2015). A comparison of measures to screen for social, emotional, and behavioral risk. *School Psychology Quarterly, 30,* 184–196. doi:10.1037/spq0000085.

Miller, F. G., Chafouleas, S. M., Welsh, M. E., Riley-Tillman, & T. C., Fabiano, G. A. (2019). Examining the stability of social, emotional, and behavioral risk status: Implications for screening frequency. *School Psychology, 34*(1), 43–53. https://doi.org/10.1037/spq0000252

Naser, S., Brown, J., & Verlenden, J. (2018). The utility of universal screening to guide school-based prevention initiatives: Comparison of office discipline referrals to standardized emotional and behavioral risk screening. *Contemporary School Psychology 22*, 424–434. https://doi.org/10.1007/s40688-018-0173-2

National Association of School Psychologists. (2020). Providing effective social–

  emotional and behavioral supports after COVID-19 closures: Universal screening

  and Tier 1 interventions [handout]. Author.

Pearson Education, Inc. (2019, March). *BASC 3 Publication Summary.*

  https://images.pearsonclinical.com/images/Assets/BASC-

  3/BASC3_Pub_Summary.pdf.

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C.,

  Hedden, S. L., Crosby, A. E., Visser, S. N., Schieve, L. A., Parks, S. E., Hall, J.

  E., Brody, D., Simile, C. M., Thompson, W. M., Baio, J., Avenevoli, S., Kogan,

  M. D., & Huang, L. N. (2013). Mental health surveillance among children: United

  States, 2005–2011. MMWR Surveillance Summaries, 62(Suppl. 2), 1–35.

  https://www.cdc.gov/mmwr/preview/mmwrhtml/su6202a1.htm

Raines, T. C., Dever, B. V., Kamphaus, R. W., & Roach, A. T. (2012). Universal

  screening for behavioral and emotional risk: A promising method for reducing

  disproportionate placement in special education. *The Journal of Negro Education,

  81*(3), 283-296.

Rollenhagen, J., Buckman, M. M., Oakes, W. P., & Lane, K. L. (2021). *Screening

  coordinator training manual. A guide for installing the Student Risk Screening

  Scale – Internalizing and Externalizing (SRSS-IE) in your school or district.* Ci3t.

  https://www.ci3t.org/ispring/screening_manual/

Romer, N., von der Embse, N., Eklund, K., Kilgus, S., Perales, K., Splett, J. W., Sudlo,

  S., Wheeler, D. (2020). Best practices in social, emotional, and behavioral

  screening: An implementation guide. Version 2.0. School Mental Health

Collaborative. https://smhcollaborative.org/wp-content/uploads/2019/11/universalscreening.pdf.

Sanchez, A. L., Cornacchio, D., Poznanski, B., Golik, A., Chou, T., & Comer, J. S. (2018). The effectiveness of school-based mental health services for elementary-aged children: A meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 57*(3), 153–165. https://doi.org/10.1016/j.jaac.2017.11.022

Splett, J. W., Chafouleas, S. M., & George, M. W. R. (2018). Accessing behavioral health services: Introduction to a special issue of research, policy, and practice. *School Mental Health, 10*, 91–95. https://doi.org/10.1007/s12310-018-9262-7

VanDerHeyden, A. M., Burns, M. K., & Bonifay, W. (2018). Is more screening better? The relationship between frequent screening, accirate decisions, and reading proficiency. *School Psychology Review, 47*(1), 62-82. doi: 10.17105/SPR-2017-0017.V47-1

Volpe, R. J., & Briesch, A. M. (2018). Establishing evidence-based behavioral screening practices in US schools. *School Psychology Review*, *47*(4), 396–402. DOI: 10.17105/SPR-2018-0047.V47-4.

Walker, H. M., Small, J. W., Severson, H. H., Seeley, J. R., & Feil, E. G. (2014). Multiple-gating approaches in universal screening within school and community 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. 47–75). American Psychological Association. doi:10.1037/14316-003