5-2019

Direct Behavior Observations and Teachers' Ratings of Internalizing Problem Behaviors

Lauren R. Parker

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

Part of the School Psychology Commons

The University of Southern Mississippi

Direct Behavior Observations and Teachers' Ratings of Internalizing Problem Behaviors

by

Lauren Parker

A Thesis
Submitted to the Honors College of
The University of Southern Mississippi
in Partial Fulfillment
of Honors Requirements

May 2019

Approved by:

        _____

Lauren McKinley, Ph.D., Thesis Adviser
Assistant Professor of Psychology

        _____

Joe Olmi, Ph.D., Director
School of Psychology

        _____

Ellen Weinauer, Ph.D., Dean
Honors College

Abstract

Systematic direct observation is a commonly used method in schools to compile data on students' behavior in a naturalistic setting. Internalizing problem behaviors, behaviors such as social withdrawal, somatic complaints, and anxiousness, can lead to adverse consequences and have not, to the researchers' knowledge, been systematically observed in a classroom setting. Additionally, students have never been identified for intervention through direct observation. The researchers administered the Student Internalizing Behavior Screener (SIBS), which is a 7-item, teacher-completed, Likert-type scale which measures occurrences of internalizing problem behaviors. Once scores from the SIBS were calculated, students who scored at-risk were chosen for observation along with students not-at-risk for use as comparison. Their behavior in a classroom setting was observed to determine if an at-risk score on the SIBS correlated with an elevated occurrence of internalizing problem behaviors in a naturalistic setting. The protocol developed for this project is the Internalizing Behavior Observation Protocol (IBOP), which includes an observation form developed by the researchers for the purposes of observing students who are at-risk for internalizing problem behaviors. Results showed that a small number of students did exhibit internalizing problem behaviors during direct observation. However, the level of internalizing problem behaviors exhibited by at-risk students was lower than expected. Further research could investigate other settings and behaviors.

Key words: internalizing problem behaviors, direct behavior observation, Student Internalizing Behavior Screener, Internalizing Behavior Observation Protocol

Dedication

I would like to dedicate this work to my family—Mom, Dad, Aden, and Gran: thank you for supporting me through this process, dealing with stressed-out phone calls, and reading the drafts. Most importantly, to my husband Colton, for helping me stay sane through getting married, graduating, and finishing a thesis all at the same time. I love you all.

Acknowledgements

First, I would like to acknowledge my thesis adviser, Dr. Lauren McKinley, who was incredibly wonderful to work with. Dr. McKinley was an amazing and patient mentor, always taking the time to answer all my questions and help me through every step of the process. She also introduced me to school psychology, an area I had no knowledge of previously, gave me innumerable opportunities for acquiring new knowledge and engaging in research, and pushed me to challenge myself.

Secondly, I would like to thank the faculty and students at Richton Elementary School, without whom this research would not be possible. The teachers' acceptance of me coming in and out of their classrooms is definitely appreciated. I would also like to thank Ashleigh Eaves, school psychology doctoral student, for all her help at the school and in completing IOA data.

Lastly, I would like to acknowledge the Honors College's support and guidance over the past four years. The faculty and staff in the Honors College had a significant influence in my enjoyment and success at USM.

Table of Contents

List of Tables

List of Abbreviations

SIBS	Student Internalizing Behavior Screener
IBOP	Internalizing Behavior Observation Protocol
DBR	Direct Behavior Ratings
MTSS	Multi-Tiered System of Support
IOA	Inter-Observer Agreement
SSBD	Systematic Screening for Behavior Disorders
SSBS	School Social Behavior Scale
RBPC	Revised Behavior Problem Checklist
ECBI	Eyberg Child Behavior Inventory
SRSS	Student Risk Screening Scale
BESS	Behavioral and Emotional Screening Scale
SABRS	Social and Academic Behavior Risk Screener
SPF-7	Student Protective Factors Screening Scale

Chapter One: Introduction

In recent years, there has been an increase in pressure to identify mental health issues in school-aged children, with this responsibility often falling on school systems (Neil & Smith, 2017). Given that teachers often have more time and opportunities to observe students than parents have to observe their children (Neil & Smith, 2017), teachers play an important role in identification of mental health issues. Yet, internalizing problem behaviors, which are often an indicator of mental health disorders, are under-identified by teachers and other school-based personnel (Walker, Severson, & Seeley, 2010), even though there is an estimated thirty-six percent prevalence rate of students with a mental health disorder at any one time (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). Internalizing problem behaviors include "sadness, worry, fear, social withdrawal, and somatic complaints related to psychological disorders such as depression and anxiety," (Hartman, Gresham, & Byrd, 2016, p. 1). Internalizing problem behaviors are inner-directed and are often over-controlled (Sanders, Merrell, & Cobb, 1999), meaning that individuals affected by internalizing problem behaviors attempt to over-regulate their internal states. Internalizing problem behaviors are grossly under-identified in students (Kauffman, 1999) because children with internalizing problem behaviors oftentimes fit the picture of a good student (e.g., they do not engage in disruptive behaviors). In addition, internalizing problem behaviors are sometimes difficult to observe. Unlike externalizing behaviors, which include behaviors such as hyperactivity and aggressive behavior, internalizing problem behaviors have fewer external indicators, though they do include behaviors such as depression, anxiety, withdrawal, and somatization (Merrell & Gueldner, 2010). If internalizing disorders are not identified

early, they can lead to undesirable outcomes, such as academic issues, dropping out of school, drug addiction, and even criminal behavior later in life (Lane, Parks, Kalberg, & Carter, 2007). Therefore, early identification via universal screening is considered to be increasingly important in schools.

This research project examines the relationship between internalizing problem behaviors and direct observations. It specifically addresses whether direct observation of students' internalizing problem behaviors directly correlates with their level of risk for internalizing disorders as identified by their scores on a universal screener for internalizing problem disorders. It is hypothesized that higher scores on the universal screener will correspond with higher levels of internalizing problem behaviors seen during direct behavior observations.

Chapter 2: Literature Review

Early Intervention and Multi-Tiered System of Support (MTSS)

Early identification leads to early intervention through identifying students in need of services. Early intervention, when intersecting with mental health, results in a system in which universal mental health screeners are considered a part of Tier 1 services, a part of Multi-Tiered System of Support (MTSS). Prevention of social and emotional difficulties is increasingly considered a priority of schools (Kauffman, 1999), even though the majority of youth with these issues will not receive the services they need (Durlak, 1997). In addition, efforts to prevent these disorders are often sparse and unorganized (Kauffman, 1999). In the vein of social and emotional problems, early identification of mental health problems or internalizing problem behaviors is considered a form of prevention and has broad implications for public health and education (Levitt,

Saka, Romanelli, & Hoagwood, 2007). If children with internalizing problem behaviors are not identified, numerous negative consequences can occur, including educational consequences, such as academic underachievement and placement in special education (Ringeisen, Henderson, & Hoagwood, 2003). Schools play a vital role in providing mental health services to their students (Ringeisen, Henderson, & Hoagwood, 2003), yet even when students are identified, often, many do not receive services due to a lack of resources in the schools, a lack of training of school-based personnel, and a perception that mental health services are an added burden to schools (Ringeisen, Henderson, & Hoagwood, 2003).

Despite these challenges, school-based personnel are moving toward the use of universal measures to identify children who may need additional support in order to prevent later adverse outcomes (Durlak, 1997). Rather than using office discipline referrals or special education assessment referrals (i.e., "refer-test-place") which can be time-consuming and resource-exhaustive, universal screenings have been proposed as a method to quickly and accurately identify students who are at-risk. Rather than a reactive model, using universal screeners serves as a proactive model to identify students with mental health problems and guide decision making about who needs services (Dowdy, Ritchey, & Kamphaus, 2010). Universal screenings for internalizing problem behaviors are used within a multi-tiered system of support (MTSS) model (Von der Embse et al., 2014) as a guide to determine the level of intensity of an intervention that is needed to best support students (Cook et al., 2011). Specifically, students who do not respond to Tier 1 (general behavioral and emotional support) are then placed at Tier 2 (targeted,

small-group support); students who do not respond appropriately at Tier 2 will then be given more individualized assistance at Tier 3.

The goal of Tier 1, as it relates to social-emotional learning, is to improve all students' social-emotional learning in schools. School-based personnel, including administration, school psychologists, school counselors, behavior specialists, and school social workers, are tasked with promoting a positive social-emotional learning environment (Merrell & Gueldner, 2010). Across the country, numerous laws and initiatives encouraging district- and building-wide efforts to improve mental health and to prevent crises from occurring exist in many places. For example, in Mississippi, the Anti-Bullying Law (HB 263, 2017) was passed in 2017 and defines behaviors that are considered bullying, and it also mandates that each school district implements a suicide prevention program, including required suicide prevention training for all staff members. Furthermore, another example of a Tier 1 social and emotional intervention is the Strong Kids curriculum, developed by Merrell and colleagues (Merrell, Carrizales, Feuerborn, Gueldner, & Tran, 2007). Strong Kids aims to promote positive emotional processes in students and has adaptations for pre-kindergarten through high school. The program aims to help children and adolescents both understand their feelings, identify cognitive distortions, create behavior change, and more (Merrell, et al, 2007). Overall, Tier 1 interventions for internalizing problem behaviors aim to prevent the development of internalizing issues and address them before they become a hindrance to students' wellbeing.

Tier 2 interventions aim at providing services to students who do not respond appropriately to Tier 1 interventions and who need a higher level of support. One strategy

for Tier 2 interventions with internalizing disorders that Merrell and Gueldner (2010) recommended is simply enriching Tier 1 interventions for certain students by providing extra time to review concepts, offer additional opportunities to practice new coping skills, and by teaching students one-on-one or in small groups (Merrell & Gueldner, 2010). Other general Tier 2 strategies that Merrell and Gueldner suggest include counseling sessions, educational sessions, and other "traditional" methods of intervention. Other, more specific programs can also be used, such as the Coping with Adolescent Depression program (Clark, Lewinsohn, & Hops, 1990) and the Coping Cat program, which are geared toward children with depression and anxiety, respectively (Kendall & Hedtke, 2006). Another Tier 2 intervention that has been proposed for internalizing disorders is using check-in/check-out as a way to prevent children at-risk for internalizing disorders from needing additional services. Check-in/check-out is a behavior intervention plan in which students meet with a mentor at the beginning and end of the day to review behavioral standards and receive feedback on behavior (Dart et al., 2015). Dart and colleagues (2015) found that training peers to check-in and check-out with students at-risk for internalizing problem behaviors is a more efficient means to implement this strategy.

Finally, there are several Tier 3 interventions for internalizing problem behaviors. One evidence-based intervention is to implement Tier 2 interventions on a more individualized level or one-on-one basis or to increase the frequency of Tier 2 interventions. This increase in intensity can be facilitated through providing in-school services, such as intervention with a school counselor or school psychologist, or through referring a student to outside community agencies. These external entities are suggested

for mental health referrals of students who may be experiencing severe symptoms, such as suicidal ideation or psychosis (Merrell & Gueldner, 2010). The authors also call for greater collaboration among medical professionals, such as pediatricians and psychologists, to enhance services provided to underserved children (Merrell & Gueldner, 2010).

Universal Screeners

With an increasing emphasis being placed on early identification of internalizing disorders as well as the use of a MTSS model, universal screeners are increasingly implemented in schools, though still at an overall low rate of less than two percent (Romer & McIntosh, 2005) and more so for academic purposes than for internalizing behavior problems. In an attempt to evaluate different universal screeners used to assess students at-risk for behavior problems, Severson and colleagues (2007) completed a fairly exhaustive analysis of the various emotional and behavioral universal screeners that were commonly used in schools, including the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1990), the School Social Behavior Scale (SSBS; Merrell, 1993), the Revised Behavior Problem Checklist (RBPC; Quay & Peterson, 1987), the Eyberg Child Behavior Inventory (ECBI; Eyberg & Ross, 1978), and the Drummond Student Risk Screening Scale (Drummond, 1993). Of these, significant limitations were found with each (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007).

The Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1990) is a universal screener that screens for both internalizing and externalizing problems; its purpose is to identify students who are at-risk for emotional and behavioral disorders. The SSBD did have positive reliability findings for both internalizing and

externalizing problems, and it also distinguished between children with and without an emotional-behavioral disorder. The SSBD was also relatively accurate, at 84%, with classifying subjects in various target groups, but the internalizing problems group was only correctly classified 64% of the time (Walker et al., 1994). However, the SSBD was considered by Severson and colleagues to be too long and expensive for schools to use and lacked long-term predictive validity (Severson, et al, 2007).

The School Social Behavior Scale, or the SSBS, was developed by Kenneth Merrell in 1993. Its main purpose is to screen for social competence and antisocial behavior via teacher and peer evaluations. Social competence subscales include interpersonal skills, self-management skills, and academic skills, while antisocial behavior subscales include hostile-irritable, antisocial-aggressive, and disruptive-demanding. The SSBS was found to have high internal consistency, moderate test-retest reliability, and significant inter-rater reliability, test validity, and construct validity. The SSBS did identify correctly seventy percent of students with a behavioral disorder (Severson, et al, 2007). However, since the SSBS was developed primarily to evaluate social competence and antisocial behavior, it is inadequate to assess for specific internalizing problem behaviors (Merrell, 1993). In addition, Severson and colleagues confirmed this finding (Severson, et al, 2007); the SSBS was found not to effectively identify students with internalizing problem behaviors (Severson, et al, 2007), though it did identify correctly seventy percent of students with a behavioral disorder. Of importance, Merrell asserted that both rating scales and systematic direct observation are most effective when combined to evaluate students' behavior.

The Student Risk Screening Scale (SRSS; Drummond, 1993) is a common universal screener in schools and has been found to have satisfactory internal consistency, test-retest stability, social validity, and predictive validity. It also has been found to have discriminative validity in dividing lower risk individuals from those with moderate to high risk (Lane, Bruhn, Eisner, & Kalberg, 2010). However, the SRSS was primarily intended to be used to assess antisocial behavior (Drummond, 1993) and was not intended to screen for internalizing behavior (Lane et al., 2010), making it a less than ideal option for screening for internalizing behavior problems.

The Revised Behavior Problem Checklist, or the RBPC, is used widely in schools and clinical settings (Hogan, Quay, Vaughn, & Shapiro, 1989) and seems promising. It is completed by a teacher or another adult familiar with the child (Hogan et al, 1989) and can be used for students ages five to eighteen (Severson., et al., 2007). It boasts a wide variety of psychometric scales, including those for Conduct Disorder, Socialized Aggression, Attention Problems-Immaturity, Anxiety-Withdrawal, Motor Excess, and Psychotic Behavior (Hogan et al, 1989). Long and colleagues demonstrated that specific scales can be chosen to represent internalizing and externalizing problems, in their case, the Anxiety-Withdrawal scale (internalizing) and Conduct Disorder scale (externalizing) (Long, Slater, Forehand, & Fauber, 1988). Although Long and colleagues demonstrated that specific subscales can be utilized to screen for specific factors, the RBPC's main purpose is more holistic in nature and is meant to screen for various behavioral disorders (Hogan, et al., 1989) and is not meant to screen exclusively for internalizing problem behaviors. The RBPC also lacks representative national norms (Severson, et al, 2007),

8

and therefore, its validity is questionable. Furthermore, its length of six subscales

prohibits its widespread use as a Tier 1 universal screener.

Similar to the SRSS, the Eyberg Child Behavior Inventory, or the ECBI, primarily

assesses disruptive behavior, leaving out significant internalizing factors (Eyberg & Ross,

1978). The ECBI consists of interviewing parents followed by direct observation in class.

Although the problem behaviors were thoroughly identified over a two-year period

(Eyberg & Ross, 1978), almost all identified behaviors had a physical or oppositional

component, which is much less applicable to internalizing disorders. The ECBI's

requirement of having a professional observe the child in class is, as well, a hindrance to

using the ECBI as a universal screener.

Severson, et al. (2007) determined the SSBD to be the most effective screener, as

it was concluded to have adequate standardization and norms while also being

inexpensive. The SSBD can be used for internalizing and externalizing problems and has

been widely used to assess students; however, it is time-consuming and may be too

cumbersome for many schools with limited resources (Lane et al., 2010).

Since Severson and colleagues completed a review of the commonly used

universal screeners that were available in 2007, three additional universal screeners have

been developed. The Behavioral and Emotional Screening Scale (BESS) is a part of the

Behavior Assessment System for Children-2 (BASC-2) that consists of student-

completed, teacher-completed, and parent-completed portions to assess for mental health

issues among children in schools across several behavioral and emotional categories

(Kamphaus & Reynolds, 2007). These categories include inattention/hyperactivity,

internalizing problem behaviors, school problems, and personal adjustments (Dowdy, et

al., 2011). The BESS has been found to have adequate test-retest reliability and convergent validity when assessed in comparison to several other rating scales, including Conner's Rating Skills and the Children's Depression Inventory (Kamphaus & Reynolds, 2007). However, despite the sound psychometric properties and fairly wide use of the BESS, it is not the best measure of internalizing problem behaviors, as more research is needed to better clarify which internalizing problem behaviors are being screened for and the level of risk being assessed (Dever, Dowdy, Raines, & Carnazzo, 2015).

Kilgus and colleagues developed and validated a universal screener called the Social and Academic Behavior Risk Screener (SABRS) to help fill a void they found in existing screeners, which they considered to be either too lengthy, costly, or lacking in assessment of key behaviors (Kilgus, Chafouleas, & Riley-Tillman, 2013). The SABRS addresses various areas of difficulties students face, including social, academic, internalizing, and externalizing problems but only assessed social and academic behavior. Although the Social and Academic Behavior Scales were found to be internally consistent, the behaviors the researchers included in the scales primarily reflected behaviors associated with externalizing disorders, such as aggression and disruptive behavior (Von der Embse et al., 2014). Although internalizing problem behaviors were addressed, they were not the primary component. Therefore, with the lack of brief, validated, and cost-effective universal screeners available to screen specifically for internalizing problem behaviors, a gap in the literature is apparent.

A novel approach to assessing students' social and emotional wellbeing involves appraising children's protective factors that may help predict students' developmental outcomes. Morrison (2015) developed the Student Protective Factors Screening Scale

(SPF-7) as a method to assess student's protective factors, including their competence and sense of purpose, social skills, respect for others, relatedness to peer groups, engagement in school, connectedness with teachers, and family support. The SPF-7 was found to have adequate test-retest reliability and predictive validity, and it also identified the approximate percentage of students expected to have social and behavioral problems (6.4% to 25.9%); however, it did not have adequate inter-rater reliability. The SPF-7 represents a preventative approach to behavioral and emotional problems as a Tier 1, public health approach to internalizing problem behaviors in schools instead of a reactive approach. However, with the lack of inter-rater reliability and the fact that this preventative approach is extremely new, the SPF-7 lacks the validity to represent the most effective universal screener.

The need for a different universal screener, without limitations such as expense, length, or a lack of validity or reliability, while still effectively screening for social and emotional problems, led Cook and colleagues to develop the Student Internalizing Behavior Screener (Cook et al., 2011). The Student Internalizing Behavior Screener, henceforth referred to as SIBS, is used to assess students' internalizing behavior symptoms. It is considered beneficial because it is brief, time-effective, and free, yet has broad implications (Hartman, Gresham, & Byrd, 2016). The Student Internalizing Behavior Screener was developed and validated in order to most accurately identify students with internalizing behavior patterns. It was modeled after Drummond's Student Risk Screening Scale (Drummond, 1994). Internal validity and temporal stability were found to be acceptable, and at a cutoff score of 8, the SIBS correctly identified 86% of students with an internalizing disorder (Cook et al., 2011). Hartman and colleagues

(2016) found the SIBS to demonstrate internal consistency and test-retest reliability. It was also found that the SIBS identified the estimated 15% of students who would need Tier 2 or Tier 3 supports (Hartman, Gresham, & Byrd, 2016). Hartman and colleagues also assert that universal screeners such as the SIBS and SEBS (Student Externalizing Behavior Screener, Cook, Volpe, & Gresham, 2012) could help alleviate the under identification of students with a diagnosis of EBD.

Direct Behavior Observation

While universal screeners are certainly useful tools for identifying the intensity of need for several students simultaneously, direct observation is also an evidence-based tool that can be used to measure students' behavior in schools (Hintze, Volpe, & Shapiro, 2002). In fact, the Individuals with Disabilities Education Act (IDEA) stipulates a school observation be completed when evaluating students with social or behavioral problems (Nock & Kurtz, 2005), and it has even been called the "gold standard of behavioral assessment measures," (Riley-Tillman, Chafouleas, Sassu, Chanese, & Glazer, 2008, p. 136). Direct observation is completed by school psychologists or other qualified personnel to measure specific behavior in a naturalistic setting (Nock & Kurtz, 2005). Direct observation can be conducted using a variety of methods of data collection. One variable in direct observation is the temporal recording that is used. For example, frequency or event recording is used when the number of times the behavior transpires is counted over a specific time period; duration recording measures the length of time the behavior occurs; and latency recording marks the length of time between a stimulus and onset of behavior (Hintze et al., 2002). Similarly, when observing students for a specific amount of time, the observer can use whole-interval, partial-interval, or momentary time

sampling recording. Of these, momentary time sampling is usually the most accurate and prone to smaller measurement errors (Alvero, Struss & Rappaport, 2008).

Direct Behavior Ratings

Much research has been done to develop Direct Behavior Ratings, or DBRs. Seen as a compromise between systematic direct observation and direct rating scales (Chafouleas, Riley-Tillman, & Christ, 2009), DBRs are unique in that they measure behavior in close temporal proximity to the observed behavior, the rater is familiar with the student, and the target behavior is reasonably discernable (Chafouleas et al., 2009). DBRs have also been suggested as a Tier 1 universal screener for problems such as social behavior (Chafouleas et al., 2009), academic engagement, and disruptive behavior (Kilgus, Riley-Tillman, Chafouleas, Christ, & Welsh, 2013). Furthermore, DBRs have been suggested as a progress monitoring tool due to their brevity and ease of completion (Riley-Tillman et al., 2008).

DBRs have many different applications, including applications for students with internalizing problem behaviors. One example of the use of DBRs with students with internalizing problem behaviors is to use DBRs as a progress monitoring tool for students with academic anxiety, specifically test anxiety (Von der Embse et al., 2014). Academic anxiety includes anxiety surrounding a particular event (i.e., a test; Zeidner, 1998) or academic subjects (i.e., math; Hembree, 1990). Test anxiety, specifically, can result in lower test grades and increased stress, among other consequences (Hembree, 1988). Von der Embse and colleagues (2014) noted a lack of regular evaluations to assess the efficacy of interventions for internalizing problem behaviors. Even though DBRs have been found to be valid and reliable as a formative behavior assessment (Chafouleas,

Riley-Tillman, & Christ, 2009) they have been used very infrequently to assess internalizing problem behaviors (Von der Embse et al., 2014). To address this gap in the literature, Von der Embse and colleagues evaluated the use of DBRs as an assessment tool with test anxiety. Results indicated that DBRs were an accurate measure of test anxiety, as DBRs in the study had sufficient concurrent validity with the Test Anxiety Inventory (1980). The authors suggest using DBRs as a Tier II measure in the future (Von der Embse et al., 2014).

Riley-Tillman and colleagues (2008) found that both DBRs and direct observation were found to have acceptable adequacy in identifying disruptive behavior in children, and indeed, direct observation and DBRs are often seen as cooperative and compatible with each other. However, some research has shown that there are several attributes of direct observation that make it preferable to DBRs. Though direct observation is less efficient than DBRs (Chafouleas, Riley-Tillman, & Christ, 2009), it is free from many potential issues that are inherent in DBRs and behavior ratings, such as rater biases and halo effects (Cohen & Kasen, 1999). Direct observation also allows clinicians to determine functions of behavior, unlike DBRs (Hanley, Iwata, & McCord, 2003). Furthermore, direct observation is usually completed in a school, which is a convenient opportunity to assess students in different settings and situations (Nock & Kurtz, 2005).

Through a thorough review of the literature, it is not apparent that any research has been completed concerning the systematic direct observation of internalizing problem behaviors of children, particularly in school settings. Although Von der Embse and colleagues (2014) completed a study using DBRs as a measure of test anxiety, and DBRs have been proposed as a type of universal screener for social and emotional problems

(Kilgus, Riley-Tillman, Chafouleas, Christ, & Welsh, 2014), to the researcher's knowledge, the use of systematic direct observation to measure internalizing problem behaviors has not been empirically examined. This is of interest given the fact that internalizing problem behaviors do, typically, have several corresponding observable behaviors, including symptoms of depression (i.e., loss of interest, crying), symptoms of anxiety (i.e., tense, easily embarrassed), social withdrawal (i.e., shy, separated from others), and somatic complains (i.e., stomachaches and headaches, exhausted; Merrell & Gueldner, 2010). Given the strengths of direct observation, it is surprising that this type of research has not been investigated yet, especially given the high rate of internalizing problem behaviors in children.

The aim of the current study is to measure the internalizing problem behaviors of students by directly observing their internalizing problem behaviors in a classroom setting. It is hypothesized that by examining students' internalizing behaviors by recording their observable behavior, internalizing problem behaviors will be made apparent. By choosing students who have scored at-risk on the Student Internalizing Behavior Screener as well as other students who did not score at-risk for the purposes of comparison, the students who scored at-risk should, therefore, have accompanying internalizing behavior problems.

Therefore, the research questions for this study are as follows:

1. Will direct observation of students' internalizing problem behaviors directly correlate with their level of risk for internalizing disorders as identified by their scores on the Student Internalizing Behavior Screener?

2. Is direct observation of internalizing problem behaviors, using the Internalizing Behavior Observation Protocol, a valid measure when compared with students' scores on the Student Internalizing Behavior Screener?

Chapter Three: Methodology

Participants

For this study, eighteen students from an elementary school in the Southeastern United States participated. Students were identified and chosen from kindergarten, first, second, third, fourth, and sixth grades. In total, the researcher randomly selected nine students who scored at-risk (minimum score of 4) to observe and randomly selected nine students who had a score of 0 in order to serve as a basis of comparison. Either two or four students from one to two classrooms in each grade (kindergarten, first, second, third, fourth, and sixth) were chosen in order to make completing observations more convenient.

Setting

Students were identified and observed at one rural elementary school in the Southeastern United States. Observations were conducted during academic engaged time, typically during the morning hours. Specifically, observations were conducted during individual seat work, whole group instruction, and small group work.

Procedures

Prior to the study, Institutional Review Board, or IRB, approval, was obtained. To receive IRB approval, the primary researcher assessed the risks to participants and determined there to be negligible risk involved. Consent was gained from the participating elementary school, and the staff were informed of the procedures in a

faculty meeting. Researchers at a university in the Southeastern United States established

a partnership with a local elementary school as part of their school psychology

department. At this school, the university began administering the SIBS, training teachers

how to use it. The SIBS was regularly administered three times throughout the school

year, once in the fall semester, once prior to winter break, and once in the spring

semester. For the purposes of this study, the scores from the fall administration of the

SIBS were used. Selected students whose scores fell above the cutoff score were

observed using the Internalizing Behavior Observation Protocol (IBOP), which has been

proposed as an observational procedure for individuals with internalizing problem

behaviors. Additionally, an equal number of students who did not score at-risk on the

IBOP were observed as a basis for comparison. After completing the observations, the

researchers compared the SIBS score with the frequency or duration of each internalizing

problem behavior and analyzed whether or not they both indicated the student to be at-

risk. Then, the researchers analyzed the total score, comparing the percentage of the

observation interval the student engaged in the behavior compared with the total score.

The thematic associations between responses and these criteria indicated an elevated risk.

The levels of risk for the direct observation procedures began at 4, which is the same

score on the SIBS that first indicates risk (Hartman et al., 2017).

Materials

     Student Internalizing Behavior Screener (SIBS). The SIBS is a universal screener

aimed at identifying internalizing problem behaviors in students. It is a teacher-completed

form that consists of seven items that teachers rate on a four-point Likert scale, from

never (0), rarely (1), sometimes (2), and frequently (3). The descriptors by which teachers

rate students include: "nervous or fearful, bullied by peers, spends time alone, clings to adults, withdrawn, seems sad or unhappy, complains about being sick or hurt," (Cook et al., 2011, p. 74). Cook and colleagues (2011) initially developed the SIBS and found internal consistency to be acceptable, with Cronbach's alpha scores of .81 and .79 in the fall and winter screenings (Cook et al., 2011), and a test-retest coefficient of .74 was considered acceptable. The SIBS was tested alongside the Teacher Report Form Internalizing Scale (TRF; Achenbach & Rescorla, 2001) and the Student Risk Screening Scale (SRSS; Drummond, 1994). The SIBS was found to strongly correlate with the TRF with a score of .82 and moderately with the SRSS with a score of 0.41. When compared with the TRF Internalizing Scale as the criterion for a cut-off score, it was found that a cutoff score of 8 correctly identified 86% of students who would have been considered at-risk on the TRF; therefore, 8 was determined to be the best cutoff score, dividing students who are at-risk from those who are not. In addition, according to Hartman and colleagues, a score in the range of 4 to 7 indicates "on the radar," (Hartman et al., 2017, p. 109). Considering these two studies, for this study, in order to develop a tiered system of risk, a score of 0 to 3 was considered no risk; a score of 4 to 7 was considered moderate risk; and a score of 8 and above was considered at-risk. All of the participants in this study were classified as "on the radar" in this study. With the increasing emphasis on early intervention, the researchers determined that "on the radar" necessitated direct observation.

   Internalizing Behavior Observation Protocol (IBOP). The Internalizing Behavior Observation Protocol (IBOP) was developed by the researchers following extensive inquiry into previous research completed on internalizing problem behaviors. Some

behaviors selected for (withdrawal and clinging) were recorded using momentary time sampling given that these behaviors are best measured by duration, while some behaviors (sadness, somatization, worry) were measured using partial interval recording as these behaviors are best measured by frequency. For each student, three observations were completed over the course of several weeks, with each observation period lasting fifteen minutes and each interval lasting fifteen seconds. Observations were completed during individual seat work, whole group instruction, and small group work in order to give the researchers the opportunity to observe the students during academically engaged periods.

Internalizing problem behaviors were selected by the researchers as they correlated to items on the SIBS. As the SIBS includes the items "nervous or fearful, bullied by peers, spends time alone, clings to adults, withdrawn, seems sad or unhappy, complains about being sick or hurt," (Cook et al., 2011, p. 74), the researchers chose to observe withdrawal, sadness, somatization, clinging, and worry. Bullied by peers was excluded due to difficulty in observing that behavior during instructional time, and spends time alone was included with withdrawal. Therefore, for the IBOP, nervous or fearful corresponds to worry; clings to adults corresponds to clinging; withdrawn corresponds to withdrawal; seems sad or unhappy corresponds to sadness; and complains about being sick or hurt corresponds to somatization.

Operational definitions of the internalizing behaviors are as follows:

1. Withdrawal (duration): Moving from a location appropriate to the classroom activity/academic task to one that is not appropriate for academic task/activity, but more isolated from others; remaining in isolated location. No attempts at

social initiation with peers or adults while in isolation; may ignore or briefly

respond to requests for attention.

2. Clinging (duration): Reluctance to leave others (e.g., parents, teachers)

    manifested by student keeping extremely close physical proximity to an

    individual, including physical contact to said individual. Does not include

    appropriate behavior with similar physical component (e.g., peer tutoring).

3. Sadness (frequency): Verbal expression of grief, unhappiness, self-

    deprecation, catastrophizing, or pessimism.

4. Somatization (frequency): Verbal expression of complaints regarding health,

    sickness, or pain.

5. Worry (frequency): Verbal expression of phobias, fear, or nervousness

    (general or specific).

Interobserver Agreement (IOA):

A secondary researcher, a doctoral student in a school psychology program, was

trained on the IBOP procedures. This researcher then conducted twenty percent of the

observations in conjunction with the primary researcher. In Table 1, "S" refers to the

student and his or her corresponding number, while "O" refers to the observation number.

The primary researcher and secondary researcher had 100% agreement between all

eleven intervals that they observed concurrently. This was considered an adequate

percentage to conclude that the observations were valid. In particular, 100% interobserver

agreement was calculated for Student 11, Observation 1 in which the behavior of sadness

was observed in two of the intervals.

Table 1: IOA Results

Student	Observer #1	Observer #2	IOA Percentage
S6/O1	0/60	0/60	100%
S7/O1	0/60	0/60	100%
S8/O1	0/60	0/60	100%
S9/O1	0/60	0/60	100%
S10/O1	0/60	0/60	100%
S11/O1	2/60	2/60	100%
S12/O1	0/60	0/60	100%
S14/O3	0/60	0/60	100%
S16/O1	0/60	0/60	100%
S17/O1	0/60	0/60	100%
S18/O1	0/60	0/60	100%

Chapter Four: Results

SIBS Results

The SIBS was completed by all teachers at the elementary school approximately one month into the school year, and students' scores were reported. Nine students whose scores indicated moderate risk (score of 4 to 7) were randomly selected for observation. No students in the school scored an 8 or above. Nine students who scored a 0 on the SIBS (not at risk) were chosen for observation as well. These students' homeroom classes corresponded to the homeroom classes of the moderate risk students selected to make observations more convenient for the researchers. These students' scores on the SIBS are reported here. Most students received a score across a range of behaviors, while three

students only exhibited two types of behavior, according to their SIBS scores. In Table 2, even student numbers correspond to moderate risk students, while odd numbers correspond to not at risk students.

Table 2: SIBS Results

Student	*Nervous or Fearful*	*Bullied by Peers*	*Spends Time Alone*	*Clings to Adults*	*Withdrawn*	*Seems Sad or Unhappy*	*Complains about Being Sick or Hurt*	**Total**
1.	0	0	0	0	0	0	0	0
2.	3	0	0	2	0	0	0	5
3.	0	0	0	0	0	0	0	0
4.	1	0	0	1	0	1	2	5
5.	0	0	0	0	0	0	0	0
6.	0	0	1	1	0	1	2	5
7.	0	0	0	0	0	0	0	0
8.	0	0	0	0	1	0	3	4
9.	0	0	0	0	0	0	0	0
10.	0	0	0	1	0	3	3	7
11.	0	0	0	0	0	0	0	0
12.	2	2	0	0	0	2	0	6
13.	0	0	0	0	0	0	0	0
14.	0	0	0	0	3	1	0	4
15.	0	0	0	0	0	0	0	0
16.	0	0	1	0	1	2	0	4
17.	0	0	0	0	0	0	0	0
18.	1	0	0	3	0	3	0	7

IBOP Results

After selecting the eighteen total students for observation, the researchers completed three, fifteen-minute long intervals of the students on three separate days. The table below gives the average percent of each behavior seen for each student over the three observations. The observations in which internalizing behaviors were seen included

Student 11, Observation 1; Student 6, Observation 2; and Student 6, Observation 3.

Overall, fewer behaviors were seen than were expected. Student 11 fell into the not-at-risk category (score of 0), while Student 6 fell into the moderate risk category (score of 5). For at least nine students (nine students observed scored moderate risk on the SIBS), the researcher expected to see elevated levels of internalizing problem behaviors during the direct behavior observations. Except for three sessions, behaviors that fit the operational definitions were not seen, though a number of behaviors that may have indicated distress were seen and are later discussed.

Table 3: IBOP Results

Student	Withdrawal	Clinging	Sadness	Somatization	Worry	TOTAL IBOP score
1	0%	0%	0%	0%	0%	0%
2	0%	0%	0%	0%	0%	0%
3	0%	0%	0%	0%	0%	0%
4	0%	0%	0%	0%	0%	0%
5	0%	0%	0%	0%	0%	0%
6	3.89%	0%	0%	0%	0%	3.89%
7	0%	0%	0%	0%	0%	0%
8	0%	0%	0%	0%	0%	0%
9	0%	0%	0%	0%	0%	0%
10	0%	0%	0%	0%	0%	0%
11	0%	0%	1.11%	0%	0%	1.11%
12	0%	0%	0%	0%	0%	0%
13	0%	0%	0%	0%	0%	0%

14	0%	0%	0%	0%	0%	0%
15	0%	0%	0%	0%	0%	0%
16	0%	0%	0%	0%	0%	0%
17	0%	0%	0%	0%	0%	0%
18	0%	0%	0%	0%	0%	0%

Percent intervals engaged in behavior across three observation sessions per each student.

Item-to-Item Analysis: SIBS and IBOP

Assigning values for scores with both the IBOP and SIBS was completed so that levels of risk could be more easily identified in order to more efficiently interpret results. Since relatively low levels of behaviors were observed, the IBOP values were as follows: 0-9 of 180 intervals (0-5%) received a value of 1; 10-18 of 180 intervals (6-10%) received a value of 2; and 19 or more out of 180 intervals (11%+) received a value of 3. A value of 1 corresponded to a low or zero risk; a value of 2 corresponded to a medium risk; and a value of 3 corresponded to at-risk. For the SIBS, a score of 0-3 received a value of 1 (no risk); a score of 4-7 received a value of 2 (moderate risk); and a score of 8 and above received a value of 3 (at-risk). These values of SIBS scores align with Hartman and colleagues' categories (2017).

These values will continue to be used when considering level of risk on the SIBS and IBOP. After consulting the scores students received on the SIBS, it was determined that 9 out of 9 students not at-risk on the SIBS were not at-risk on the IBOP (value of 1 on both the SIBS and IBOP). One student (Student 11) did exhibit sadness for a short period during one observation, though the number of behaviors exhibited did not warrant falling into the category of medium risk or at-risk. Furthermore, 9 out of 9 students

24

identified as at-risk on the SIBS were found not to be at-risk according to the IBOP results. Although one student did exhibit withdrawal during two different observations, the percentage was not high enough to warrant being given a value of 2 or 3 to indicate risk. While the results were not as expected, possible reasons for this discrepancy are discussed in Chapter 4.

When analyzing each item on the SIBS and IBOP, it was found that of the 4 students with indicated risk for "nervous or fearful" on the SIBS, none of these students exhibited worry on the IBOP. Of the 5 students with indicated risk for "clings to adults" on the SIBS, none of these students exhibited clinging on the IBOP. Of the 3 students with indicated risk for "withdrawn" on the SIBS, none of these students exhibited withdrawal on the IBOP. In fact, the one student who did exhibit withdrawal behavior on the IBOP was not identified as at-risk on the SIBS. Of the 7 students identified with indicated risk for "seems sad or unhappy" on the SIBS, none of these students exhibited sadness on the IBOP. The one student who did exhibit sadness was a student with no risk across all behaviors; this student was one of the 9 students used as comparison. Of the 4 students with indicated risk for "complains about being sick or hurt" on the SIBS, none of these students exhibited somatization on the IBOP. In Table 4, the number in the parentheses in the total columns reflects the value assigned that score, as defined by the values mentioned previously.

Table 4: Item-to-Item Comparison of SIBS and IBOP Results

	SIBS	IBOP (%)	SIBS	IBOP (%)	SIBS	IBOP (%)	SIBS	IBOP (%)	SIBS	IBOP (%)	TOTAL SIBS	TOTAL IBOP
Student	Withdrawn	Withdrawal	Clings to Adults	Clinging	Seems Sad or Unhappy	Sadness	Complains About Being Sick or Hurt	Somatization	Nervous or Fearful	Worry		
1	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
2	0	0	2	0	0	0	0	0	3	0	5 (2)	0% (1)
3	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
4	0	0	1	0	1	0	2	0	1	0	5 (2)	0% (1)
5	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
6	0	3.89	1	0	1	0	2	0	0	0	5 (2)	3.89% (1)
7	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
8	1	0	0	0	0	0	3	0	0	0	4 (2)	0% (1)
9	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
10	0	0	1	0	3	0	3	0	0	0	7 (2)	0% (1)
11	0	0	0	0	0	1.11	0	0	0	0	0 (1)	1.11% (1)
12	0	0	0	0	2	0	0	0	2	0	6 (2)	0% (1)
13	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
14	3	0	0	0	1	0	0	0	0	0	4 (2)	0% (1)
15	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
16	1	0	0	0	2	0	0	0	0	0	4 (2)	0% (1)
17	0	0	0	0	0	0	0	0	0	0	0 (1)	0% (1)
18	0	0	3	0	3	0	0	0	1	0	7 (2)	0% (1)

Note: Since two behaviors from the SIBS were excluded from the study, the totals listed in this table are not always reflective of the behaviors shown in the table. The totals displayed are the totals calculated after adding the scores from all seven original SIBS behaviors.

Overall, the hypotheses of this thesis were not supported. The first research question, "Will direct observation of students' internalizing problem behaviors directly correlate with their level of risk for internalizing disorders as identified by their scores on the Student Internalizing Behavior Screener?" was not supported. The majority of students with a moderate risk score on the SIBS did not exhibit any internalizing problem behaviors during the direct behavior observations. In addition, all 9 students with a moderate risk score failed to be given a value above 1 on their total IBOP score, which would have indicated risk. Additionally, the second research question, "Is direct observation of internalizing problem behaviors, using the Internalizing Behavior Observation Protocol, a valid measure when compared with students' scores on the Student Internalizing Behavior Screener?" was not supported due to the same reasons. Additional research is needed to confirm whether the IBOP is a valid measure for observing internalizing problem behaviors.

Chapter Five: Discussion

Limitations and Future Direction

In this study, there were several factors that may have limited the number of behaviors observed during the observations. One of the most significant limitations was that of the nine participants chosen for observation due to their elevated scores on the SIBS, all were in the "on the radar" category, according to Hartman and colleagues (2017). No students at the elementary school scored above a 7 on the SIBS, meaning that no students were at-risk, according to the cutoff score of 8 determined by Cook and colleagues (2011). If students with a higher score on the SIBS were included in this study, there might have been higher levels of behaviors observed on the IBOP. In

addition, the two major studies validating the SIBS (Cook et al., 2011 and Hartman et al., 2017) determined two different sets of criteria for an optimal SIBS cut off score. Cook and colleagues (2011) found that 8 was the ideal cutoff score, while Hartman and colleagues (2017) created the "at-risk" category consisting of a score of 4 to 7. Therefore, there is some variability in interpretation guidelines as to which cutoff score has the most validity, although it can be assumed that a higher score would correspond to more behaviors seen in the classroom.

Although the SIBS has had several studies that have ensured its validity (Cook et al., 2011 and Hartman et al., 2017), the IBOP is a tool developed specifically for this study and therefore has no prior details on validity and reliability. Most importantly, a cutoff score is not yet available for the IBOP, meaning that there is no definition for what level of observed behavior should be considered "typical". For example, is a low level of withdrawal, sadness, etc. normal in a classroom setting, even with a score of 0 on the SIBS, or should students who score a 0 on the SIBS also exhibit no behaviors in the classroom? Considering that one student who exhibited behavior in an observation was not at-risk, it could be hypothesized that seeing a low level of these behaviors in not-at-risk students should not be considered problematic.

Several environmental factors were limitations in this study. For one, three of the behaviors (sadness, somatization, and worry) had to be expressed verbally by the student in order to be counted as an instance of the behavior. In some classrooms, it was impossible for the observer to hear the student, either due to classroom noise or the observer having to sit at a farther than ideal distance from the student. In addition, the researchers noticed as the study progressed that small group time was the optimal time to

complete observations, as it gave the students ample opportunity to exhibit the verbal-only behaviors mentioned before. The researchers made efforts to observe students during small group time or when students were more likely to exhibit verbal behaviors, such as during whole group instruction.

A suggestion for future study in this area would be to expand the behaviors on the IBOP (or other direct observation tool). Specifically, in several observations, the researchers observed at-risk students exhibiting gestures that suggested distress, such as covering ears with hands and holding hands up. In addition, students were observed exhibiting off-task and inattentive behavior during several observations. The inclusion of off-task behavior could be a topic of further study, and it could be helpful to determine whether off-task behavior could be indicative of internalizing problem behaviors. Including other behaviors in future studies could help researchers understand what other behaviors are and are not characteristic of internalizing problem behaviors.

Another future consideration includes the time of day, class, teacher, and other factors that may influence students' behavior patterns. All observations for this study were completed during instructional time, and no definitive correlations were seen between internalizing problem behaviors and subject or teacher. However, during non-instructional time, some students might exhibit a higher level of behaviors. In addition, some students might exhibit behavior in one specific subject or in one specific teacher's class. Classroom environment, including other students in the class, could be hypothesized to play a large role in students' internalizing problem behaviors. Abry and colleagues found that higher levels of classroom-level adversity, which include risk factors such as family problems, poor nutrition, and other problems that affect students in

the classroom, correlated with higher levels of internalizing and externalizing behaviors (Abry et al., 2017). Therefore, classroom factors such as subject and teacher as well as external problems that affect students in the classroom (classroom-level adversity) could contribute to students' internalizing problem behaviors and should be further studied.

Several additional topics could be of further study. For example, one teacher commented that students are more likely to exhibit internalizing problem behaviors during testing, specifically standardized testing. Therefore, completing a similar study during and around the time of standardized testing (if it is allowed) might yield a higher level of exhibited behaviors. This type of research might indeed yield higher levels of internalizing problem behaviors, as there have been many studies completed showing heightened test anxiety among students, especially surrounding high-stakes testing (i.e., Segool, Carlson, Goforth, Von der Embse, & Barterian, 2013). Finally, the connection between internalizing and externalizing behaviors should be further studied. Although a number of screening instruments have been developed to screen for both of these (i.e., SIBS, Cook et al., 2011; SABRS, Kilgus, et al 2013), and several screeners combine both internalizing and externalizing behaviors (i.e., SABRS, Kilgus, et al 2013; SSBD Walker & Severson, 1990), the connection between the two has not been extensively studied. In this study, students at-risk would occasionally exhibit externalizing behaviors, such as talking out of turn. Since externalizing behaviors include disruptive behaviors (Sanders et al., 1999), then talking inappropriately could be considered an externalizing behavior. A suggestion for further study would be to observe both internalizing and externalizing behaviors systematically in the same sample of at-risk students to see if a correlation exists between the two.

References

Abry, T., Bryce, C.I., Swanson, J., Bradley, R.H., Fabes, R.A. & Corwyn, R.F. (2017). Classroom-level adversity: Associations with children's internalizing and externalizing behaviors across elementary school. *Developmental Psychology 55*(3), 497-519.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles: An integrated system of multi-informant assessment.* Burlington, WT: Research Center for Children, Youth, and Families.

Alvero, A. M., Struss, K., & Rappaport, E. (2008). Measuring Safety Performance. *Journal of Organizational Behavior Management, 27*(4), 1-28.

Anthoine, E., Moret, L., Regnault, A., Sébille, V., & Hardouin, J. (2014). Sample size used to validate a scale: A review of publications on newly-developed patient reported outcomes measures. *Health and Quality of Life Outcomes, 12*(1). doi:10.1186/s12955-014-0176-2

Anti-Bullying Law, MS-H.R. 263, Regular Session (2017).

Chafouleas, S. M., Riley-Tillman, T. C., & Christ, T. J. (2009). Direct Behavior Rating (DBR): an emerging method for assessing social behavior within a Tiered intervention system. *Assessment for Effective Intervention, 34*(4), 195-200. doi:10.1177/1534508409340391

Clarke, G., Lewinsohn, P., & Hops, H. (1990). Introduction. In *leader's manual for adolescent groups adolescent coping with depression course* (pp. 2-7). Portland, OR: Kaiser Permanente.

Cohen, P., & Kasen, S. (1999). The context of assessment: Culture, race, and

socioeconomic status as influences on the assessment of children. In D. Shaffer,

C. P. Lucas, & J. E. Richters (Eds.), *Diagnostic assessment in child and*

adolescent psychopathology (pp. 299-318). New York, NY, US: Guilford Press.

Cook, C. R., Rasetshwane, K. B., Truelson, E., Grant, S., Dart, E. H., Collins, T. A., &

Sprague, J. (2011). Student internalizing behavior screener: examination of

reliability, validation, and classification accuracy. *Assessment for Effective*

Intervention, 36(2), 71-79. doi:10.1037/t21577-000

Cook, C., Volpe, R., & Gresham, F. M. (2012). *Technical adequacy, classification*

accuracy, and social validity of the Student Externalizing Behavior Screener.

Unpublished manuscript.

Costello, E.J., Mustillo, S., Erkanli, A., Keller, G., & Angold, A. (2003). Prevalence and

development of psychiatric disorders in childhood and adolescence. *Archives of*

General Psychiatry, 60, 837-844.

Dart, E. H., Furlow, C. M., Collins, T. A., Brewer, E., Gresham, F. M., & Chenier, K. H.

(2015). Peer-mediated check-in/check-out for students at-risk for internalizing

disorders. *School Psychology Quarterly, 30*(2), 229-243. doi:10.1037/spq0000092

Dever, B. V., Dowdy, E., Raines, T. C., & Carnazzo, K. (2015). Stability and change of

behavioral and emotional screening scores. *Psychology in the Schools, 52*(6), 618-

629. doi:10.1002/pits.21825

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: a

population-based approach to inform and monitor children's mental health

needs. *School Mental Health, 2*(4), 166-176. doi:10.1007/s12310-010-9036-3

Drummond, T. (1993). *The student risk screening scale (SRSS).* Grants Pass, OR:

    Josephine County Mental Health Program.

Durlak, J. A. (1997). Prevention of behavioral and social problems. In *Successful*

    prevention programs for children and adolescents (pp. 27-53). New York, NY:

    Springer-Verlag New York.

Von der Embse, N. P., Scott, E., & Kilgus, S. P. (2015). Sensitivity to change and

    concurrent validity of direct behavior ratings for academic anxiety. *School*

    Psychology Quarterly, 30(2), 244-259. doi:10.1037/spq0000083

Eyberg, S. M., & Ross, A. W. (1978). Assessment of child behavior problems: The

    validation of a new inventory. *Journal of Clinical Child Psychology, 7*(2), 113-

    116. doi:10.1080/15374417809532835

Hair, J.E., Anderson, R.E., Tatham, R.L., Black, W.C. (1995). *Multivariate data*

    analysis: with readings. Englewood Cliffs, NJ: Prentice Hall.

Hanley, G. P., Iwata, B. A., & Mccord, B. E. (2003). Functional analysis of problem

    behavior: A review. *Journal of Applied Behavior Analysis, 36*(2), 147-185.

    doi:10.1901/jaba.2003.36-147

Hartman, K., Gresham, F. M., & Byrd, S. (2017). Student Internalizing and Externalizing

    Behavior Screeners: Evidence for Reliability, Validity, and Usability in

    Elementary Schools. *Behavioral Disorders, 42*(3), 108-118.

    doi:10.1177/0198742916688656

Hembree, R. (1988). Correlates, causes, effects, and treatment of text anxiety. *Review of*

    Educational Research, 58, 47-77.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21,* 33-46.

Hintze, J. M., Volpe, R. J., & Shapiro, E. S. (2002). Best practices in the systematic direct observation of student behavior. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (993-1006). Bethesda, MD: National Association of School Psychologists.

Hogan, A. E., Quay, H. C., Vaughn, S., & Shapiro, S. K. (1989). Revised behavior problem checklist: stability, prevalence, and incidence of behavior problems in kindergarten and first-grade children. *Psychological Assessment, 1*(2), 103-111. doi:10.1037//1040-3590.1.2.103

Kauffman, J. M. (1999). How We Prevent the Prevention of Emotional and Behavioral Difficulties in Education. *Exceptional Children, 65*(4), 448-468. doi:10.4135/9781848608146.n27

Kendall, P. C., & Hedtke, K. A. (2006). *Cognitive–behavioral therapy for anxious children: Therapist manual* (3rd ed.). Llanfair, PA: Workbook Publishing.

Kilgus, S. P., Chafouleas, S. M., & Riley-Tillman, T. C. (2013). Development and initial validation of the social and academic behavior risk screener for elementary grades. *School Psychology Quarterly, 28*(3), 210-226. doi:10.1037/spq0000024

Kilgus, S. P., Riley-Tillman, T. C., Chafouleas, S. M., Christ, T. J., & Welsh, M. E. (2014). Direct behavior rating as a school-based behavior universal screener: Replication across sites. *Journal of School Psychology, 52*(1), 63-82. doi:10.1016/j.jsp.2013.11.002

Lane, K. L., Bruhn, A. L., Eisner, S. L., & Kalberg, J. R. (2010). Score reliability and

    validity of the student risk screening scale: a psychometrically sound, feasible tool

    for use in urban middle schools. *Journal of Emotional and Behavioral Disorders,*

    18(4), 211-224. doi:10.1177/1063426609349733

Lane, K. L., Parks, R. J., Kalberg, J. R., & Carter, E. W. (2007). Systematic Screening at

    the Middle School Level. *Journal of Emotional and Behavioral Disorders, 15*(4),

    209-222. doi:10.1177/10634266070150040301

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(2), 163-191. doi:10.1016/j.jsp.2006.11.005

Long, N., Slater, S., Forehand, R., & Fauber, R. (1988). Continued high or reduced

    interparental conflict following divorce: Relation to young adolescent adjustment.

    Journal of Consulting and Clinical Psychology, 56(3), 467-469.

Merrell, K. W. (1993). Using behavior rating scales to assess social skills and antisocial

    behavior in school settings: development of the school social behavior scales.

    School Psychology Review, 22(1), 115-133.

Merrell, K. W., Carrizales, D., Feuerborn, L., Gueldner, B. A., & Tran, O. K. (2007).

    Strong kids: A social and emotional learning curriculum. Baltimore: Brookes.

Merrell, K. W., & Gueldner, B. A. (2010). Preventive Interventions for students with

    internalizing disorders: Effective strategies for promoting mental health in

    schools. In M. R. Shinn & H. M. Walker (Authors), *Interventions for achievement*

    and behavior problems in a three-Tier model including RTI. Bethesda, MD:

    National Association of School Psychologists.

Morrison, J. Q. (2015). Technical adequacy of the student protective factors screening

    scale (spf-7) as a universal screening tool. *Psychology, 6*, 817-832.

    doi:10.4236/psych.2015.67081

Neil, L., & Smith, M. (2017). Teachers' recognition of anxiety and somatic symptoms in

    their pupils. *Psychology in the Schools, 54*(9), 1176-1188. doi:10.1002/pits.22055

Nock, M. K., & Kurtz, S. M. (2005). Direct behavioral observation in school settings:

    Bringing science to practice. *Cognitive and Behavioral Practice, 12*(3), 359-370.

    doi:10.1016/s1077-7229(05)80058-6

Riley-Tillman, T. C., Chafouleas, S. M., Sassu, K. A., Chanese, J. A., & Glazer, A. D.

    (2008). Examining the agreement of direct behavior ratings and systematic direct

    observation data for on-task and disruptive behavior. *Journal of Positive Behavior*

    Interventions, 10(2), 136-143. doi:10.1177/1098300707312542

Ringeisen, H., Henderson, K., & Hoagwood, K. (2003). Context Matters: Schools and the

    "Research to Practice Gap" in Children's Mental Health. *School Psychology*

    Review, 32(2), 153-168.

Romer, D., & McIntosh, M. (2005). The roles and perspectives of school mental health

    professionals in promoting adolescent mental health. In D. L. Evans, E. B. Foa, R.

    E. Gur, H. Hendin, C. P. O'Brien, M. E. Seligman, et al. (Eds.), *Treating and*

    preventing adolescent mental health disorders: What we know and what we don't

    know: A research agenda for improving the mental health of our youth (pp. 597-

    615). New York, NY: Oxford University Press.

Sanders, D.E., Merrell, K.W., & Cobb, H.C. (1999). Internalizing symptoms and affect of

    children with emotional and behavioral disorders: A comparative study with an

    urban African American sample. *Psychology in the Schools, 36*(3), 187-197.

Segool, N.K., Carlson, J.S., Goforth, A.N., Von de Von der Embse, N., Barterian, J.A.

    (2013). Heightened test anxiety among young children: Elementary school

    students' anxious responses to high-stakes testing. *Psychology in the Schools,*

    50(5), 489-499.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M.

    (2007). Proactive, early screening to detect behaviorally at-risk students: Issues,

    approaches, emerging innovations, and professional practices. *Journal of School*

    Psychology, 45(2), 193-223. doi:10.1016/j.jsp.2006.11.003

Spielberger, C. D. (1980). *Preliminary professional manual for the Test Anxiety*

    Inventory (TAI). Palo Alto, CA: Consulting Psychologists Press.

Walker, H. M., Severson, H. H., Nicholson, F., Kehle, T., Jenson, W. R., & Clark, E.

    (1994). Replication of the Systematic Screening for Behavior Disorders (SSBD)

    Procedure for the Identification of At-Risk Children. *Journal of Emotional and*

    Behavioral Disorders ,2(2), 66-77. doi:10.1177/106342669400200201

Zeidner, M. (1998). *Test anxiety: The state of the art.* New York, NY: Plenum Press.

Appendix A: Internalizing Behavior Observation Protocol

Student's Name: _____

Observer's Name: _____

Date: _____

Observation (circle one):   1   2   3

Momentary time sampling:
Withdrawal: Moving from a location appropriate to the classroom activity/academic task to one that is not appropriate for academic task/activity, but more isolated from others; remaining in isolated location. No attempts at social initiation with peers or adults while in isolation; may ignore or briefly respond to requests for attention.
Clinging: Reluctance to leave others (e.g., parents, teachers) manifested by student keeping extremely close physical proximity to an individual, including physical contact to said individual. Does not include appropriate behavior with similar physical component (e.g., peer tutoring).

Partial interval recording:
Sadness: Verbal expression of grief, unhappiness, self-deprecation, catastrophizing, or pessimism.
Somatization: Verbal expression of complaints regarding health, sickness, or pain.
Worry: Verbal expression of phobias, fear, or nervousness (general or specific).

Momentary	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Withdrawal																
Crying																
Clinging																
Partial																
Sadness																
Somatization																
Worry																

Momentary	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Withdrawal																
Crying																
Clinging																
Partial																
Sadness																
Somatization																
Worry																

Momentary	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Withdrawal																
Crying																
Clinging																
Partial																
Sadness																
Somatization																
Worry																

Momentary	49	50	51	52	53	54	55	56	57	58	59	60
Withdrawal												
Crying												
Clinging												
Partial												
Sadness												
Somatization												
Worry												

Appendix B: Student Internalizing Behavior Screener

Directions: Please rate each student on each behavior using the following scale:
0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Frequently
For each student, write the number that corresponds to the frequency rating in each cell.

Student Name	Nervous or fearful	Bullied by peers	Spends time alone	Clings to adults	Withdrawn	Seems sad or unhappy	Complains about being sick or hurt

Appendix C: IRB Approval

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

INSTITUTIONAL REVIEW BOARD
118 College Drive #5147 | Hattiesburg, MS 39406-0001
Phone: 601.266.5997 | Fax: 601.266.4377 | www.usm.edu/research/institutional.review.board

NOTICE OF COMMITTEE ACTION

The project 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 (45 CFR Part 46), and university guidelines to ensure adherence to the following criteria:

- The risks to subjects are minimized.
- The risks to subjects are 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 regarding risks to subjects must be reported immediately, but not later than 10 days following the event. This should be reported to the IRB Office via the "Adverse Effect Report Form".
- If approved, the maximum period of approval is limited to twelve months.
  Projects that exceed this period must submit an application for renewal or continuation.

PROTOCOL NUMBER: 18091001
PROJECT TITLE: Direct Behavior Observations and Teachers' Ratings of Internalizing Problem Behaviors
PROJECT TYPE: Honor's Thesis Project
RESEARCHER(S): Lauren Parker
COLLEGE/DIVISION: College of Education and Human Sciences
SCHOOL: Psychology
FUNDING AGENCY/SPONSOR: N/A
IRB COMMITTEE ACTION: Expedited Review Approval
PERIOD OF APPROVAL: 10/3/2018 to 10/3/2019

Edward L. Goshorn, Ph.D.
Institutional Review Board