

Screening in School-Wide Positive Behavior Supports:
Methodological Comparisons

by

Morgan Hall, M. A.

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2012 by the
Graduate Supervisory Committee:

Linda Caterino, Co-Chair
Sarup Mathur, Co-Chair
Jenefer Husman

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

Many schools have adopted programming designed to promote students' behavioral aptitude. A specific type of programming with this focus is School Wide Positive Behavior Supports (SWPBS), which combines positive behavior techniques with a system wide problem solving model. Aspects of this model are still being developed in the research community, including assessment techniques which aid the decision making process. Tools for screening entire student populations are examples of such assessment interests. Although screening tools which have been described as "empirically validated" and "cost effective" have been around since at least 1991, they have yet to become standard practice (Lane, Gresham, & O'Shaughnessy 2002). The lack of widespread implementation to date raises questions regarding their ecological validity and actual cost-effectiveness, leaving the development of useful tools for screening an ongoing project for many researchers. It may be beneficial for educators to expand the range of measurement to include tools which measure the symptoms at the root of the problematic behaviors. Lane, Gresham, and O'Shaughnessy (2002) note the possibility that factors from within a student, including those that are cognitive in nature, may influence not only his or her academic performance, but also aspects of behavior. A line of logic follows wherein measurement of those factors may aid the early identification of students at risk for developing disorders with related symptoms. The validity and practicality of various tools available for screening in SWPBS were investigated.

DEDICATION

This work is dedicated to Leslie Croy and Jeff Hair for their unconditional support during its production.

ACKNOWLEDGMENTS

It is an honor to acknowledge Dr. Linda Caterino, Dr. Sarup Mathur, and Dr. Jenefer Husman for the guidance and patience that made this dissertation possible. I would also like to thank Meridith Swenson and Sara Banegas for the time and energy put forth during this project.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	v
CHAPTER	
1 INTRODUCTION.....	1
2 LITERATURE REVIEW	3
Purpose of the Study	29
3 METHOD	33
4 RESULTS	36
5 DISCUSSION	43
Conclusion	53
REFERENCES	55
APPENDIX A	60
APPENDIX B	62

LIST OF TABLES

Table		Page
1.	<i>Correlations of Screening Instruments and Daily Behavior Outcome.....</i>	<i>54</i>
2.	<i>Correlations of Screening Instruments and Daily Behavior Outcome.....</i>	<i>54</i>

Chapter 1

Introduction

Social, emotional, and behavioral competencies are critical to a child's development and as such are important to educators. The development of these skills has been found to be necessary for future adjustment (Dishion & Patterson, 1999). Additionally, these skills seem to be complexly intertwined with academic success (Kalberg, Lane, & Menzies, 2010). For these reasons, many schools have adopted programming designed to foster the growth of students' social, emotional, and behavioral aptitudes. One type of programming is School-Wide Positive Behavior Supports (SWPBS), which integrates positive behavioral techniques with a system-wide problem solving model. The research community is still developing aspects of this model, including assessment techniques that aid the decision-making process. Assessment-related research interests include developing tools for screening entire student populations to determine their risk for not developing these socio-emotional skills appropriately, which is important for determining which students need extra services.

Universal screening is a process of collecting assessment data for a given population with the intent of identifying risk factors predictive of a specified disorder. It is an important aspect of SWPBS because it contributes to early identification of children at risk for behavioral maladjustment. Early identification is necessary for early intervention, and early intervention is valuable because it offers an opportunity to teach adaptive behaviors before problematic ones become engrained in a student's behavioral pattern (Forness, 2000). Several methods for

universal screening are available. One method involves using data that already exists within most schools, such as office disciplinary referrals. An alternative is to develop tools specific to social, emotional, and behavioral screening procedures. This method is exemplified by the use of brief behavior rating scales. A third method is to adapt measurement techniques from other areas of psychology and education for the purpose of universal screening in SWPBS.

Psychometric properties valuable in most assessment techniques, such as reliability and validity, are important characteristics of screening tools as well. With screening, though, elements of practicality also become a great concern because of the quantity of resources needed when focusing on an entire school; especially considering limitations on the availability of time and money. Thus, the cost-effectiveness and efficiency with which the instrument can be administered are also important qualities of a universal screening tool. The goal of this study is to compare and contrast various universal screening methods.

Chapter 2

Literature Review

Behavior in Schools

In their service to students, schools and educational systems do more than provide academic instruction. They are also charged with developing appropriate behaviors in their students, and behavior management comprises a major responsibility of teachers and administrators. This is important because research has shown that students with behavioral concerns are far more likely to be deficient in basic academic skills as compared to their peers without such difficulties (Reid, Gonzalez, Nordness, Trout, & Epstein, 2004) and are at much greater risk of school failure (Kauffman & Landrum, 2009; Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005).

The relationship between effective instruction and classroom management makes student behavior a concern for educators. Behaviors that are norm- or rule-violating, disruptive, and challenging for schools to manage are problematic in an educational setting. This makes behavior management strategies, particularly those that are positive and preventative in nature, some of the best methods of supporting effective instruction (Sugai & Horner, 2008). Additionally, appropriate instruction can be a strong behavior management tool, partly because it increases student engagement in academic tasks and decreases problem behavior. This adds to the entangled relationship between academic instruction and behavior management (Sugai & Horner, 2008; Walker, 2010).

The existence of a relationship between behavior, namely maladaptive externalizing behavior, and academic underachievement is evident (Lane, Grasham, & O'Shaughnessy, 2002; Kalberg, Lane, & Menzies, 2010). However, the particular nature of the relationship is less clear and a number of different models have been proposed to account for it (Hinshaw, 1992). One possible model supposes that students with a lack of academic skills will exhibit problems with externalizing behavior as a coping mechanism enabling them to avoid school work that may be too difficult for them. Alternatively, another proposed relationship asserts that children with externalizing behavior problems develop deficiencies in academic skills as the result of reduced academic engagement due to the behavior itself or consequences for the behavior (Kalberg, Lane, & Menzies, 2010). It is also possible that the relationship is bidirectional with academic skills deficits and maladaptive sets of behaviors mutually inciting one another. Finally, the possibility exists that other factors from the environment or within the child influence academic achievement and behavior problems (Hinshaw, 1992; Lane, Grasham, & O'Shaughnessy, 2002). The specific nature of the relationship between academic skill development and externalizing behavior problems is a complicated issue, but it may be important at the individual student level for its instrumentality in the guidance of intervention planning.

Sugai and his colleagues (Sugai & Horner, 2008; Sadler & Sugai, 2009) presume that when schools can promote academic engagement by creating a supportive, constructive culture, they become more effective learning

environments. Nurturing behavioral development positively influences academic development, and further illustrates the relationship between learning and behavior in the classroom. In his 2001 report on youth violence, the Surgeon General commented on the status of behavioral problems in youth and provided several recommendations including eradicating antisocial peer networks, raising academic performance, building positive school climates, and implementing primary prevention efforts (Satcher, 2001). Too often, though, there is a lag between onset of problematic behaviors or mental health issues and appropriate service delivery, with only a fraction of the students in need of social or behavioral services actually receiving them (Forness, et al., 2000; Walker, 2010).

Lane, Gresham, and O'Shaughnessy (2002) suggest that a continuum of intervention efforts would be preferable to late or missed opportunities for promoting development of adaptive behaviors. This continuum would begin early with a focus on prevention and then shift to intervention efforts as needed. For example, it would begin in preschool and transition over the course of a student's education through the 12th grade. Initial intervention efforts would focus on the prevention of emotional and behavioral problems. Subsequent intervention efforts would be directed at remediation, amelioration, and finally, accommodation in the later years of secondary education (Lane, Gresham, & O'Shaughnessy 2002).

Proactive Prevention Strategies

Use and rationale

Positive behavior support (PBS) practices, which have a focus on prevention, are valuable because they benefit both educational outcomes and social behavior, which are mutually reliant on each other (Sugai & Horner, 2008). Preventative efforts are those that seek to decrease the development, likelihood of future occurrences, and aggravation of emotional and behavioral problems. This approach discourages punishment-oriented reactive practices in an attempt to avoid their associated negative side effects. The shortcomings of punishment-oriented reactive practices include lack of effectiveness, increases in antisocial behavior, more coercive interactions among students and staff, and reductions in academic achievement and displays of social behaviors. Primary prevention efforts, on the other hand, can include screening all children within a school setting in an attempt to identify students at risk for developing maladaptive behaviors. They also include implementation of programming designed to teach adaptive skills in social and behavioral domains (Forness, et al., 2000). Examples of types of possible primary prevention provided by Forness and colleagues (2000) include parent training programs, as well as teacher-led interventions.

PBS minimizes the effects of negative behaviors and the costs of dealing with them. Ideally, many children who are initially at risk will never go on to develop problematic behaviors that interfere with their education (Forness, et al., 2000; Sugai & Horner, 2008). Other advantages include maximizing protective factors, and being more economical overall because they prevent costs associated

with later discipline (Forness et al., 2000). One of the primary benefits of a prevention model is the avoidance of the negative effects associated with labeling students, which has been reported to be aversive to the parents of students identified as emotionally disturbed (ED) (Lane, et al., 2002; Forness et al., 2000). Another argument for the use of preventative approaches is that interventions aimed at individual students are better implemented within school-wide supportive contexts (Gresham, 2005; Sadler & Sugai, 2009; Sugai & Horner, 2008). Moreover, along with other important educational mandates more academic in nature, IDEA (2004) calls for the implementation of Positive Behavior Interventions and Supports (PBIS).

School wide positive behavior supports

School Wide Positive Behavior Supports (SWPBS) has positive primary prevention efforts at its foundation, and also includes a problem solving system with secondary and tertiary interventions. It is an approach to school- or classroom-wide behavior management which combines a primary prevention framework with applied behavioral modification techniques. It has its roots in Behaviorism, the theoretical orientation associated with psychologists such as B.F. Skinner (1953), relying most on positive reinforcement of desired behaviors (i.e., presenting a student with something pleasing after performing a desirable or required behavior). Reinforcement, by nature, increases the likelihood of a given behavior occurring in the future and is more effective at teaching appropriate behaviors than its punitive counterpart. Punishment, such as the presentation of an aversive consequence or removal of something desirable, can also be effective in

decreasing the frequency of an undesired behavior. However, it does not offer practice in more appropriate behaviors, leaving room for the development of other maladaptive behaviors. The use of reinforcement techniques in SWPBS allows educators to teach students more adaptive behaviors which can serve many of the same functions associated with negative behaviors (Vollmer & Iwata, 1992). This makes negative behaviors obsolete and ultimately creates a more positive atmosphere (Vollmer & Iwata, 1992). Data-based decision making, human behavior sciences, validated practices, and procedures for systems change are also incorporated into SWPBS (Sugai & Horner, 2008; Sadler & Sugai, 2009).

The Model

SWPBS emphasizes effective behavioral interventions at the systemic and individual level with a goal of enhancing social and learning outcomes (Sugai & Horner, 2008). It is a 3 tiered model for promoting adaptive behavior within the school setting and simultaneously reducing problematic ones. There is an emphasis on primary prevention at the first tier, followed by secondary and tertiary interventions when needed. Primary efforts involve several components (Kalberg, Lane, & Menzies, 2010). First, social and behavioral expectations written in a positive manner are developed by educators. More specifically, the expectations are framed by stating what the students *should* do as opposed to what they *should not* do (i.e., the word “no” is avoided in phrases). Goals are expressed in a positive manner. These expectations are taught to all students, and students are reinforced by every adult within the school setting when they meet expectations (Kalberg, Lane, & Menzies, 2010).

At the first tier in SWPBS, also referred to as the universal tier, practices showing evidence of increasing positive behaviors among students, are implemented for the entire school or class. All students, staff, and school settings are involved. Another component at this level is population screening for behavioral and/or emotional problems. The universal screening is for the purpose of early identification, as early intervention is important to prevent maladaptive behaviors from becoming engrained in a student's pattern of behaviors (Forness et al., 2000). Accordingly, universal screening procedures can contribute to improved behavioral health in schools (Albers, Glove, & Kratochwill, 2007).

Beyond the universal tier, additional intervention endeavors take place at tier 2 and are targeted at smaller groups, typically 5-10% of the school population. If those interventions are not effective to a satisfactory extent, increasingly targeted and intense interventions are applied on an individual basis in tier 3. As the tiers are ascended, interventions become more targeted and intensive. Thus, the interventions become more idiosyncratic to the child, are more specific to the problem at hand, and can be more time consuming. Furthermore, policy makers and researchers alike stress the importance of using research-based interventions, and interventions or practices lacking empirical evidence are avoided (Lane, Gresham, & O'Shaughnessy, 2002; Sugai & Horner, 2008). It is also important to note that the translation of research-based practices to actual students, teachers, and classrooms must be well thought-out with appropriately developed adaptations. In order to provide more effective services to children and their

families, a wide range of interventions that are effective, efficient, and relevant to academic and behavioral domains will be needed.

In addition to research-based interventions, the use of data plays an integral role in monitoring SWPBS systems. Data are used to assess the efficacy of specific program interventions in a school and then to make decisions as to which programs to continue. Data analysis is also used in progress monitoring to keep track of student improvement and to determine which students may need more intensive interventions in tiers 2 & 3 (Lane, et al., 2002; Sugai & Horner, 2008). Similar to interventions, data collection will become increasingly targeted, idiosyncratic to the child, and time intensive as the tiers are progressed. Data can and should be collected through various means at the different levels. For example, screening data collected at the universal level might come from brief behavior rating scales, tier 2 data might take the form of daily behavior report cards, with functional behavioral analyses being used at the 3rd tier (Lane, et al., 2002; Kalberg, Lane, & Menzies, 2010).

SWPBS assumes that behavior is teachable, predictable, affected by contextual and physiological factors, and can be manipulated by the environment (U. S. Department of Education, 2010). The implementation of SWPBS has been shown to reduce behavior problems and increase academic success in schools (Sugai & Horner, 2008; Walker, 2010). Furthermore, SWPBS that includes early intervention can be beneficial because it reduces the need to label children as disabled before directing efforts to their behavioral needs. In addition to being beneficial to the future welfare of the child, SWPBS practices bolster academic

instruction and intervention by furnishing students with the skills to engage in their learning materials appropriately (Kalberg, Lane, & Menzies, 2010). However, the efficacy of SWPBS is reliant on the fidelity of its implementation. The quality of the programming selected (i.e., evidence based practices), the validity of the data used in decision making, and the consistency with which the staff are able to adhere to programming and curricula are important characteristics of successful SWPBS. The fidelity of a SWPBS system is compromised when the quality of these characteristics is unreliable, making it important for educators to know which intervention programs and data collection tools are both valid and practical. Sugai & Horner (2008) suggest that resources from education, public health, child and family welfare, juvenile justice, and psychological services should be combined to support school based mental health via a comprehensive system, so it is possible that related fields will need to be called upon to make advancements in SWPBS.

Current Status and Limitations

SWPBS systems have many parallels with academic problem solving systems such as Response to Intervention (RTI). The rise of RTI, a legally authorized method for determining eligibility for special education services for learning disabilities, has illuminated the lack of progress in the area of empirically validated services for children with, or at-risk for developing E/BD (Briesch, Chafouleas, & Riley-Tillman, 2010; Gresham et al., 2010). A distinction between the two terms, RTI and SWPBS, should be made. Response to intervention includes the use of systematically more intense tier-related interventions and the

related data on which such decisions are based with the possibility of ultimately leading to determination of eligibility for special education services when a student fails to respond to academic interventions. SWPBS also includes tiered levels of intervention and data based decision making, but does not necessarily qualify students for special education. Some use terms like “Response to Intervention for Social, Emotional and Behavior Domains” or a shortened version thereof when SWPBS data is used for such determinations (Fellers et al., 2010; Saeki, 2010). However, there are several aspects of SWPBS which need improvement and further research before this is fully appropriate. These areas which remain underdeveloped limit the ability of SWPBS approaches to serve an RTI function for emotional and behavioral disorders.

For example, in academic response to intervention general outcome measures which can be administered efficiently have been developed to screen for children who are likely to demonstrate academic difficulties in the future. Consequently, struggling students can receive early interventions to target areas with inadequate progress. Reading is likely the area with the most development in problem solving systems thus far, with basic reading skills having the most developed probes. Thus, it serves as a consummate example of early identification and intervention against which other areas of pupil services can be compared. In RTI for reading, measures of oral reading fluency show a comfortable level of validity evidence in their ability to predict future problems in reading overall (Reschly, 2009). Although measures of oral reading fluency do not represent the construct of reading ability in its entirety, they correlate well with measures of

other aspects of reading, such as reading comprehension, making it a widely agreed upon general outcome measure (National Institute of Child Health and Human Development, 2000). Another advantageous aspect of oral reading fluency measures, in addition to the efficiency with which they can be administered, is that they can be utilized early in the school year before a teacher has had ample time to ascertain which students lack adequate reading skills resulting in opportunities for even earlier intervention. Few such measures have been proposed for use in the emotional and behavioral domains, but it has been noted that systematic screeners at the universal level are an important aspect of merging SWPBS with RTI (Horner & Sugai, 2008). Without the ability to screen all students before problems arise, it is difficult to identify students in need of early intervention, and without early intervention, it is difficult to prevent the development of maladaptive behaviors and related psychopathologies in students at risk. In addition to screening tools, progress monitoring tools are essential to problem solving models. However, as Gresham et al. (2010) describe, there are no curriculum based measure (CBM) analogues for gauging students' responses to social, emotional, and behaviorally based interventions as of yet. Thus, further study and advancement of psychometric tools for use in SWPBS seems warranted.

The psychometrics of SWPBS

The use of SWPBS has highlighted the growing need for research with regard to data collection (Brisch, Chafouleas, & Riley-Tillman, 2010; Kalberg, Lane, & Menzies, 2010; Volpe & Gadow, 2010). Two assessment related

concerns in SWPBS are that of screening and progress monitoring. Collecting data on all students is important a) for identifying students at risk for developing an emotional or behavioral problem, and b) to inform educators regarding interventions that are appropriate for the specific risk factors at hand. Without universal screening procedures in place, opportunities for early intervention may be missed, and decisions regarding which students to target may not be empirically based. This is particularly dangerous when resources, including staff time and program funding, are limited within a school because it may lead to an uneconomical allotment of those resources (Walker, 2010). Similar concerns apply to progress-monitoring data collection. Without data reflecting the progress of students participating in tier 2 and tier 3 interventions, it is difficult to determine who is responding to the intervention and who might benefit from a higher level of service or a change of service. Therefore, universal screening and progress monitoring are integral components of SWPBS.

Progress monitoring

Progress monitoring, necessary for students at tiers 2 & 3, can take many forms. These include office disciplinary referrals (ODRs), brief behavior rating scales (BBRSs), daily behavior report cards (DBRs), and functional behavioral analyses (FBAs) (Gresham et al., 2010; McIntosh, Frank, & Spaulding, 2010; Volpe & Gadow, 2010). Each of these data gathering techniques has advantages and drawbacks (Volpe & Gadow, 2010). The target of measurement for a progress monitoring tool should reflect the intervention for which progress is being

monitored, creating a psychometric challenge considering that diverse interventions exist for a spectrum of behavioral problems.

Volpe and Gadow (2010) describe two possible methods for creating change-sensitive and relatively psychometrically sound progress monitoring tools, specifically BBRs, with the recognition that it is not feasible for teachers to use lengthy behavior report forms such as those that would typically be included in a psychoeducational evaluation [e.g., BASC-II (Reynolds & Kamphous, 2004); Teacher Report Form (Achenbach & Rescorla, 2001)]. BBR methods generally consist of 2 subtypes— nomothetic and idiographic. With nomothetic approaches, scales with sound psychometric properties are created based on large samples. In one nomothetic approach, the items with the highest factor loadings are isolated from the larger bank of questions on a scale and used for regularly occurring progress monitoring. The other nomothetic approach involves using items that are likely to be most sensitive to change. These two approaches rely heavily on statistical procedures, making their use reliant on the statistical background of the school staff. With idiographic approaches, educators can use knowledge about their individual students to create briefer rating scales. One method allows the reporter to select the items that they think are relevant to the student of interest from one or several existing rating scales. Alternatively, the responder can complete a rating scale initially, and the items rated highest as being problematic can be selected for continuous progress monitoring. The drawbacks to idiographic approaches include challenges with evaluating the psychometric properties of the instrument, and the phenomenon of regression to the mean. The latter refers to the

fact that subsequent scores may be closer to the mean, which would falsely reflect response to intervention. However, Volpe and Gadow (2010) provide support for the use of either type of shortened rating scale in progress monitoring over the use of full length rating scales that are more time intensive. These methods seem to be present in the development of screening procedures also, as screening procedures are meant to identify students not making expected progress after being exposed to universal interventions.

The use of ODRs in monitoring the progress that a student is making with a certain behavioral intervention shows some utility, but is also accompanied by drawbacks (McIntosh, Frank, & Spaulding, 2010; Pas, Bradshaw, & Mitchell, 2011). The benefit of using ODRs lies in that they are already collected by most schools. However, McIntosh, Frank, and Spaulding (2010) state that when ODRs take the form of “unstandardized incident reports” their validity can be problematic. The information provided in those instances is too inconsistent to be considered readily useful. Further, referral practices differ within and across schools. To be useful, ODRs should be standardized for behavior, location, and time. With predefined choices, ODRs become more consistent and efficient for teachers to use. However, McIntosh, Frank, and Spaulding, (2010) recommend their use only in conjunction with other data for decision making. Additionally, ODRs are essentially measures of negative behaviors and it is also necessary to measure positive behaviors when assessing intervention efficacy, which is also more consistent with the theoretical framework of SWPBS.

Universal screening

Screening students for being at risk for developing an emotional and/or behavioral disorder is another psychometrically based concern in the field of SWPBS. Burk et al. (2012) describe universal screening within SWPBS as a systematic approach intended to identify students displaying problem behaviors or demonstrating features that may place them at behavioral risk. They further note that it is intended to proactively identify students who are nonresponsive to school wide prevention and primary intervention efforts. Some screening tools currently in existence overlap with progress monitoring tools for SWPBS; however, the purpose and scope of measurement is generally different, sometimes inspiring a necessity to create different instruments. Although screening tools which have been described as “empirically validated” and “cost effective” have been around since at least 1990, they have yet to become standard practice, even in schools following SWPBS models [e.g. Social Skills Rating System (Gresham & Elliot, 1990)] (Lane, et al., 2002). The lack of widespread implementation raises questions regarding the ecological validity and actual cost-effectiveness, leaving the development of useful tools for screening purposes an ongoing project for many researchers. Also, Albers, Glover, and Kratochwill (2007) argue that the identification of assessments and outcomes of screening that are acceptable and valuable to schools is an important research endeavor.

Current methods for screening in SWPBS include ODRs and BBRs (Kamphaus, Distefano, Dowdy, Eklund, & Dunn, 2010; McIntosh, Frank, & Spaulding, 2010). Frequency and type of behavior reported on ODRs have been

found to be predictive of older students' future receipt of more ODRs (McIntosh, Frank, & Spaulding, 2010), however, they do not necessarily constitute an early screener, as it can take an entire semester or school year to garner the appropriate amount of ODR data (Tobin et al., 1996; Tobin & Sugai, 1999). Less is known about the use of ODRs as a screener in elementary school, but current evidence provided by McIntosh, Frank, and Spaulding (2010) reflects that they are most responsibly used as a secondary source of information partly due to their technical inadequacy, but also because teachers' decisions to issue them can be related to ethnic bias and other subjective decisions. BBRs seem to have had an increase in production recently to meet the needs of efficiency and psychometric adequacy of screening tools. A challenge associated with BBRs as screeners include being sensitive enough to detect a variety of emotional and/or behavior problems and concurrently brief enough to be practical for teachers to use (Kamphaus, et al., 2010). As with most forms of assessment, psychometric properties such as practicality and validity of SWPBS screeners should be maximized to make them more useful to educators. This means that they need to be as short as possible, while also being relatively reliable and ecologically legitimate. Furthermore, Kalberg, Lane, and Menzies, (2010) note that teachers unaccustomed to problem solving models such as RTI may be less motivated to participate in screening procedures, especially those that are time intensive. Teachers and other education staff face many demands on their time, making it essential that screening for academic and behavioral risk factors be as efficient as possible.

Universal screening is an important aspect of any problem solving model. It is important that tools used for that purpose continue to be developed to the extent that widespread adoption takes place so that children at risk for developing emotional or behavioral disorders are identified as early as possible to provide them with appropriate levels of interventions without delay. The current methods for screening students' risk for developing an emotional or behavioral disorder rely on measuring other people's perceptions of the students' behavior (BBRSs), or actual problematic behaviors that have occurred and been recorded (ODRs). However, it may be beneficial for educators and psychologists to expand the range of measurement to include tools which measure the symptoms at the root of the problematic behaviors. Lane, Gresham, and O'Shaughnessy (2002) note the possibility that factors from within a student, both emotional and cognitive in nature, may influence not only his or her academic performance, but also aspects of the child's behavior. A line of logic follows wherein measurement of those factors may aid the early identification of students at risk for developing disorders with related symptoms.

Internal Processes and Behavior

Although behavior is an outwardly observable process (Watson, 1930), it is influenced by the internal factors of a child which can be both cognitive and affective in nature (Eisenberg, et al., 2000). Indeed, cognitive and affective factors seemingly interact to incite the development of behavior problems. Further, various forms of self-regulation, such as attentional, emotional, and behavioral regulation have been found to be related to proneness to anger (Derryberry &

Rothbart, 1988), externalizing problem behavior and conduct disorder (Eisenberg et al., 1996), aggression (Hart, Keller, Edelstein, & Hofmann, 1998), and psychopathology (Patterson & Newman, 1993). Eisenberg and colleagues (2000) used structural equation modeling to illustrate such relationships. In their study, an “attentional control” factor was comprised of parent and teacher reports, and a behavioral regulation factor was comprised of parent reports, teacher reports, as well as a frustrating performance task requiring inhibition. The two factors were not only mutually influential, but also predictive of behavior problems, especially in children with high levels of negative emotionality. Understanding the influence of internal factors may help to inform educator’s behavior management techniques and interventions. Furthermore, Eisenberg and her colleague’s (1996; 2000) work demonstrates the use of a variety of instruments (i.e., parent report forms, teacher report forms, and tasks in which children actively engage) to measure the internal processes related to external behavior.

Cognitive Functions and Behavior

Historically, behavior has been described both independent of and in association with cognitive factors (Bandura, 1977; Watson, 1930). As previously mentioned, the reinforcement component of SWPBS has its roots in behaviorism (Skinner, 1953). However, some theorists have described behavior in relation to underlying cognitions, and some subsequent interventions designed to target skills associated with behavioral problems are more cognitive in nature in that they involve self-monitoring and metacognitive strategies (Ellis, 1962). Thus, in the endeavor to have assessment inform intervention, it may be useful to incorporate

consideration of cognitive factors, especially executive functions, in the screening of children for risk of developing a behavioral disorder.

Executive functioning

Executive functions bring behavior under the control of internal processes and permit greater goal directed action and task persistence (Barkley, 1997).

They are important to success in the classroom with respect to both learning and behavior, and impairments in these skills can lead to difficulty in the academic setting (Garcia-Berrera, Kamphaus, & Bendalos, 2011; Gioia, Isquith, Guy, & Kenworthy, 2000). They are important for the success of students because they influence one's ability to self-regulate emotions and behavior, maintain sufficient levels of attention, and problem-solve.

Executive functions are discrete cognitive abilities, and examples include skills such as, but not limited to, inhibitory control, working memory, and focusing and sustaining attention, which some believe can be integrated to form one overarching factor of executive functioning. Barkley (1997) proposes that a hierarchical formation of executive functioning exists in which inhibition creates the occasion for other executive functions to occur, namely working memory, self-regulation of affect-motivation-arousal, internalization of speech, and reconstitution (i.e. behavioral analysis and synthesis), which ultimately lead to greater goal-directed behavior with increased motor control, fluency, flexibility, and persistence. Although inhibition does not cause the subsequent executive functions to occur, without inhibition there would be no opportunity for them to

take place (Barkley, 1997). Thus, inhibition may be one of the most important executive functions for classroom success.

According to Barkley (1997), behavioral inhibition takes several forms, which include prevention of initial or dominant responses, ceasing an ongoing response, and protecting a period of delay from competing events to allow for responses appropriate for the completion of a goal-directed behavior. Tasks that necessitate delays in gratification, disruption of consequences that have been perceived as sequential, generation of novel responses, and problem solving are most taxing on executive functions such as behavioral inhibition and self-regulation. Inhibition is assessed by an individual's performance on tasks that require refraining from a response, delaying a response, terminating an ongoing response, and resisting distraction. Furthermore, tasks testing behavioral inhibition relate most closely to parent or teacher ratings of hyperactive-impulsive behavior and social competence than do tasks testing other executive functions (Barkley, 1997).

It may also be important for those concerned with children's emotional-behavioral functioning to consider that disorders of the brain structures related to executive function (i.e., the prefrontal cortex) also lead to problems regulating affect because once elicited, emotions are regulated by self-directed executive functions (Barkley, 1997; Eisenberg et al., 1996). For success in academic settings, children will need the ability to regulate motivation, drive, and arousal in order to change variables causing anger, frustration, disappointment, sadness, anxiety, or boredom (Barkley, 1997; Eisenberg et al., 1996). Thus, executive

functions appear to be related not only to behavior, but affective experiences influencing behavior.

Brain and Behavior Relationships

Cognitive factors associated with observable behaviors are associated with cortical and subcortical regions of the brain. It is the prefrontal cortex which affords individuals the ability to think about the emotional consequences of a behavior before acting, and also to consider a behavior after it has been completed (Davidson, 2000). Further, dysfunction in the orbital prefrontal cortex and its connection to the ventral-medial region of the striatum results in problems with inhibition. Injury to this area has been associated with symptoms such as emotional lability, irritability, poor judgment, antisocial behavior, distractibility, and other socially inappropriate behavior (Barkley, 1997; Hale & Fiorello, 2004). There is also evidence to show that this area of a child's brain is related to psychopathologies such as Attention Deficit Hyperactive Disorder (ADHD) when underactive, and Obsessive Compulsive Disorder (OCD) when overactive (Hale & Fiorello, 2004). Other areas of the brain showing associations with behavior, particularly self-regulation of behavior, attention, and working memory, are the ventral anterior cingulate and dorsolateral prefrontal area (Hale & Fiorello, 2004). The processes by which the frontal and subcortical areas of the brain interact to affect the behavioral and emotional functioning of a child is beyond the scope of most educators; however, the actual behavioral and emotional functioning is of concern and worth measuring. The related field of neuropsychology works to do that, with specific consideration of the neurological substrates involved.

Neuropsychological assessment and education

The neuropsychological approach is not very different from the psychoeducational assessments already taking place in schools. They both often include measures of general intellectual ability, academic achievement, and social emotional functioning among other factors (Riccio, Sullivan, & Cohen, 2010). However, the neuropsychological approach to measuring children's abilities often includes the addition of several brief tasks that aim to measure discrete cognitive functions, such as executive functions, language, memory and learning, sensorimotor, social perception, and visuospatial processing (Korkman, Kirk, & Kemp, 2007). These tasks, when considered together, illustrate the child's strengths and weaknesses and inform recommendations for intervention (Riccio, Sullivan, & Cohen, 2010). Some of the functions measured by neuropsychological tasks are associated with overall behavioral regulation, and because the tasks require the child to perform some activity, they are often referred to as "performance tasks."

Purpose of performance tasks

Performance tasks are typically designed to measure a discrete ability by having the student engage in a standardized activity requiring that ability and comparing his or her performance to that of same-aged peers. For example, tasks exist to measure executive functioning in general, as well as more specific components such as inhibition and attention. Current efforts in SWPBS for screening children for problems with emotional and behavioral functioning focus on the measurement of behaviors that have already occurred (as seen with ODRs)

and others' perceptions of a student's behavior (BBRSs). Because overt behavior is influenced by factors internal to the child such as attention regulation, behavioral regulation skills such as inhibition as well as other executive functions, and emotionality (Eisenberg, et al., 2000), tests which measure these features are an important component of assessing a child's risk for problematic behavior. Thus, it seems more proactive to measure skills, such as executive functioning skills, associated with those behaviors prior to the onset of problematic behavior ultimately leading to greater prevention.

Several tests exist for the purpose of assessing individuals' executive functioning. The NEPSY-II is one such test designed specifically for school-aged children aged 3-16 years, although not all available subtests are normed for each of those ages (Korkman & Kemp, 2007). The tasks on the NEPSY-II measure a variety of functions including those that are cognitive, academic, social, and behavioral in nature, which may be helpful in school-based intervention planning (Davis & Matthews, 2010). Normative data for the NEPSY-II were reportedly collected from 2005-2006 (Davis & Matthews, 2010). Data from the October 2003 Census were analyzed to inform recruitment of a normative sample closely resembling the U.S. population of children ages 3 to 16 years old. Stratification occurred across age, race/ethnicity, geographic region, and parent education level variables. The normative sample of the NEPSY-II consisted of 1200 children with 100 children in each of the 12 age groups. Race/ethnicity categories were White, African American, Hispanic, and Other. Children with diagnoses of Attention Deficit/Hyperactive Disorder (ADHD), Reading Disorder, Language Disorder,

Mathematics Disorder, Mental Retardation, Autistic Disorder, Traumatic Brain Injury, Deaf and Hard of Hearing, and Emotionally Disturbed were included in the normative sample (Korkman & Kemp, 2007). Inter-rater reliability of scoring ranged from 93%-99% across subtests (Davis & Matthews, 2010). Most of the subtests also showed adequate to high internal consistency with reliability coefficients above .80 (Brooks, Sherman, & Strauss, 2010).

As part of its standardization, the NEPSY-II was administered to 30 students categorized with Emotional Disturbance (ED) then matched with peers in a control group for age, sex, race/ethnicity, and parent education for comparison purposes. The results showed impairments across many domains including attention and executive functioning, memory and learning, sensorimotor, and visual spatial processing (Korkman & Kemp, 2007). From the Executive Functioning domain, subtests showing statistically significant mean differences between the ED group and the control group were Animal Sorting, Response Set, Clocks, and the Inhibition time variables. However, not each of these subtests, nor items from the subtests are, are standardized for use with all ages. It is also important to avoid an assumption that a given subtest clearly measures what is implied by its title. Many tasks include a degree of construct irrelevant variance, and may be measuring more than one discrete skill. Thus, the following subtests measure more than one specific executive function, and should therefore be considered a more general measure of executive functioning which influences behavioral regulation and goal directed behavior.

Auditory Attention Subtest

The Auditory Attention subtest of the NEPSY-I I requires that children listen to a recording of a string of words, some of which correspond to simultaneously presented visual stimuli, and requires that they respond by pointing to the visual stimuli according to a set of directions previously delivered by an examiner. It is a measure of selective and sustained attention and results in scaled scores for number of omissions, number of commissions, and number of inhibitory errors. Based on the normative sample, it is appropriate for use with children aged 5-16 years and takes 7-11 minutes to administer (Brooks, Sherman, & Strauss, 2010). This subtest was used by European researchers to identify reduced executive functions in children with very-low birth weight, and compared subtest results with parent ratings on a questionnaire of behavior and development (Lind, Haataja, Rautava, Väliäho, & Lehtonen, 2010). The study found reduced scores in the affected sample as compared with a control group, as well as a significant correlation between subtest scores and parents ratings of related behaviors.

Inhibition Subtest

The Inhibition subtest of the NEPSY-II was designed as a measure of one aspect of executive functioning. It requires rapid naming of shapes and inhibiting learned responses to provide an alternative response. Three raw scores include total time for the speeded naming trial in seconds, total time for switching trials, total errors, and number of self-corrected errors. Based on the normative sample, it is appropriate for use with children aged 5-16 years and takes 8-11 minutes to

administer (Brooks, Sherman, & Strauss, 2010). The Inhibition Subtest of the NEPSY-II was used to study boys with severe behavior problems, and the subtest was compared to the Externalizing Problem subscale of the BASC-2 and Child Behavior Checklist (CBCL) (Hirayama, 2011). The researcher found that boys with severe behavior problems scored lower than the normative population on the NEPSY-II Inhibition Combined Scaled Score. The boys in this study had slower completion times on both the inhibition and switching tasks and also made more errors compared to the normative population. However, no significant relationships were found between externalizing behavior on the rating scales and Inhibition or between aggressive behavior on the rating scales and Inhibition.

Statue Subtest

The Statue subtest requires a child to stand in one posture for 75 seconds while the examiner attempts to distract him or her as a measure of executive functioning and behavior regulation behavioral inhibition. Based on the normative sample, it is appropriate for use with children aged 3-6 years and takes three minutes to administer (Brooks, Sherman, & Strauss, 2010). The subtest was used in a study of children with William's syndrome to illustrate deficits in executive control and behavioral difficulties. The study also showed a relationship between performance based tasks of executive functioning, including the Statue subtest, and parents ratings of "dysexecutive behavior" (Gallo, 2009).

Purpose of the Study

The emotional and behavioral functioning of students is important for school systems to consider because it influences many factors associated with the successful education of students such as academic engagement and children and adolescents' future adjustment (Dishion & Patterson, 1999; Kalberg, Lane, & Menzies, 2010). When the social, emotional, and behavioral development of a student is at risk, interventions are warranted in the educational setting. Systems to make intervention more efficient have been developing for decades, but there is still room for progress with respect to efficacy and data based decision making.

Three-tiered models exist for use by educators in both the academic and emotional and behavioral domains for the purpose of systematic problem solving, and research in both domains continues to grow as the models are constantly updated to include greater breadth and depth of knowledge regarding how to service students in need within such a framework. Not surprisingly, the two domains are not mutually distinct from one another. Kalberg, Lane, and Menzies (2010) have shown evidence that without appropriate SWPBS support, academic interventions may not be utilized to their full potential. For example, if a child is having problems sustaining attention or engaging in academic material appropriately, (skills which could be addressed with behavioral interventions), efforts put forth by educators for academic interventions may be thwarted leading to the student not responding to the academic intervention fully for reasons unrelated to a learning disability. Thus, it is important to screen for behavioral difficulties in conjunction with screening for academic skills deficits. Further,

Kalberg and colleagues (2010) suggest that academic and behavioral components of instruction should work in tandem. So once children have been screened for both academic and behavioral risk factors, they should receive the respective interventions from both domains as needed to optimize success.

Screening provides educators with the opportunity to identify students who may have previously gone unidentified for an undesirable amount of time and can also provide information regarding the type of interventions that may be warranted. Universal screening increases the likelihood that risk factors will be identified early. Then, interventions can be employed before problematic behavior patterns are fully developed, and negative consequences such as psychopathology and delinquency are avoided (Dishion & Patterson, 1999). However, for data from screening procedures to inform effective intervention, the data itself must be valid and meaningful. Thus, the most useful screening tools would provide reliably accurate information about students, and have the ability to be administered as early as possible in a school year.

Currently, research and evidence for best practices in problem solving models for the provision of social, emotional, and behavioral areas lags behind what is available in academic domains (Gresham, et al., 2010). Work is needed to establish not only efficacious interventions, but also proper assessment instruments including screeners and progress monitoring tools. Discovery of proper screening tools is important because of their role in identifying students not responding to the universal intervention, or in other words, at risk for developing a behavior disorder and associated pathological sequelae, therefore

warranting inclusion in a tier 2 intervention. Several categories of screening tools are developing and include parent rating scales, teacher rating scales, and tracking of ODRs. However, as educational psychologists continue to develop new instruments, further assessment of their validity and practicality is important. It is also possible that the adoption of measures, specifically performance based tasks, from the related field of neuropsychology will help fill the void in available tools for universal screening. The purpose of this study is to evaluate current tools available for screening students for emotional and behavioral problems at the universal level to determine if they need a tier 2 intervention within school-wide positive behavior intervention and support systems. This will be accomplished by addressing three research questions.

Research Question 1. The first research question is in regard to the relationship between several performance based tasks (which measure executive functioning skills) and reports of children's behavioral functioning as assessed by a behavior rating scale.

Hypothesis 1. The expectation is that some tasks will be more highly related to reports of children's behavior, with reduced functioning on performance tasks associated with higher risk of behavioral problems. A strong relationship between a performance task and behavior ratings will warrant consideration of the performance task for inclusion in research question two.

Research Question 2. The second research question considers the abilities of brief behavior rating scales and performance tasks to predict 1) teachers' reports of behavioral and emotional problems at the end of the semester and 2)

number of office disciplinary referrals at the end of the semester. The methods of screening will include parent rating scales, teacher rating scales, and a performance based task measuring cognitive based factors associated with behavioral maladjustment.

Hypothesis 2. It is expected that teachers' rating scale responses will have a more modest relationship with both their later reports and with office disciplinary referrals than the parent reports and performance tasks.

Research Question 3. The third research question concerns the utility of using a variety of screening tools in conjunction, rather than relying solely on one measure to predict emotional and behavioral problems independently.

Hypothesis 3. It is expected that the addition of multiple tools upon universal screening will enhance the ability to predict behavioral problems as determined at the end of the semester.

Chapter 3

Method

SWPBS aims to foster the development of adaptive behavior in children through systemic problem solving, data-based decision making, and tiers of intervention. The focus of this project is to identify data collection tools for use at the universal level that are both practical and show evidence of validity. The evaluation of available methods for screening children for need for intervention within SWPBS was accomplished through a two part study.

Part I.

The first research question concerns the relationship between several performance based tasks measuring cognitive and behavioral skills and ratings of children's behavioral functioning. Lane, Grasham, and O'Shaughnessy (2002) describe how factors from within a student, both emotional and cognitive in nature, may influence aspects of his or her behavior. Additionally, Riccio, Sullivan, and Cohen (2010) note that neuropsychological tasks, including performance tasks, illustrate a child's strengths and weaknesses in these domains and inform recommendations for intervention. Thus, the purpose of part one is to determine eligibility of the performance tasks for inclusion in part two of the study from an empirical standpoint as opposed to one that is solely theoretical in nature. A strong relationship between a performance task and behavior ratings will warrant consideration of the performance task for inclusion in part two, in which it will be evaluated for predictive validity as a behavior screening tool.

Participants

Participant data for part one is from an archival database of children seen in an outpatient clinic in a southwestern metropolitan area. The 60 children in the available database ranged in age from 3-16 years. They were referred to the clinic by outside sources, and tested by a clinical professional on an individual basis.

Procedure

Following approval for analysis from the Arizona State University institutional review board, archival data from a private assessment clinic in a southwestern metropolitan area was used to explore the relationships between several performance based measures of various cognitive skills, such as executive functioning, with the internalizing and externalizing scales of the Behavior Assessment Scales for Children (BASC-2) parent report form.

Materials

Performance tasks- NEPSY-II

The performance tasks are selected subtests from the NEPSY-Second Edition (NEPSY-II) Executive Functioning domain. The specific subtests for this study were chosen because of their prevalence in the available database and age appropriateness based on standardization. They included the Auditory Attention and Inhibition subtests (see Literature Review for detailed descriptions and validity information). The Statue subtest was not able to be evaluated due to limited data for children with BASC-2 PRS scores for the age group appropriate for the Statue subtest.

Behavior Assessment Scales for Children-2 (BASC-2)

The BASC-2 assesses multiple domains of child behavior (Reynolds & Kamphaus, 2004). The BASC-2 was chosen because of its common use in the field and its relation to the screening tool used in part two of the study. Five formats exist including Parent Rating Scale (PRS), Teacher Rating Scale (TRS), Self-Report of Personality (SRP), as well as the Structured Developmental History (SDH) and Student Observation System (SOS). The archival data for this study comes from a private clinic with most referrals originating from parents, thus the PRS is most widely available for data analysis in part one of this study. It measures both adaptive skills and problem behaviors (Reynolds & Kamphaus, 2004). The PRS form uses a four-choice Likert scale with responses of Never, Sometimes, Often, and Almost Always to items on the composites of Externalizing Problems, Internalizing Problems, Adaptive Skills, and Activities of Daily Living. Hyperactivity, Aggression, and Conduct Problems are subscales of the Externalizing Problems Composite. On the Internalizing composite are the Anxiety, Depression, and Somatization subscales. There is also a Behavioral Symptoms Index (BSI) available which reflects a broad composite of overall problem behaviors. The PRS was normed on 4,800 children closely matched to the 2001 *Current Population Survey* (Tan, 2007) regarding sex, mother's education level, race/ethnicity, geographic region, and special education classification variables. The PRS-C, which is appropriate for children aged 6-11 years, was used in this study. The PRS-C showed a correlation of .46 with the TRS-C in the standardization sample.

Chapter 4

Results

A multiple regression analysis was conducted to assess the abilities of the two performance tasks to predict parents' perceptions of behavior problems as measured by *T*- scores on the BASC-2 Externalizing scale. The linear combination of performance tasks was not significantly related to the BASC-2 scores, $F(2, 29) = 2.72, p = .08$. The correlation of .40 between the Auditory Attention subtest only and the BASC-2 Externalizing scale was statistically significant ($p = .01$).

Part II

The second research question addresses the abilities of three individual methods of screening student behavior to predict teachers' records of behavioral problems at the end of the semester and 2) number of office disciplinary referrals at the end of the semester.

Participants

The sample included 66 first grade students from 4 classrooms in a southwestern metropolitan city. The students ranged in age from 6 years 0 months to 7 years 4 months at the beginning of the semester. The sample included 32 boys and 34 girls. Parents, teachers, and school administration also participated by providing children's behavior outcome data. The school from which the sample was drawn has a Title I designation.

Materials

Performance task

Several factors were considered when choosing the performance task to use in Part II of the study. First, the correlation between the subtests and the Externalizing subscale of the BASC-2 from Part I of the study was considered. Second, NEPSY-II subtests' abilities to predict externalizing behavior problems in other literature was taken into account. Third, the mean differences in scores between children with emotional and behavioral disorders versus a control group as provided in the NEPSY-II technical manual was considered. The Auditory Attention and Statue subtests showed abilities to predict ratings of externalizing behavior in other literature, and the Inhibition subtest showed the most substantial difference in mean scores between children with emotional and behavior disorders and control group children as reported in the NEPSY-II Technical Manual (Kemp & Korkman, 2007). Auditory Attention, which was the most convincing measure based on two out of the three criteria, was administered to each student participating in part two of the study on an individual basis. Because the Statue subtest was not able to be evaluated in Part I of the study, it was also included in the materials for Part II of the study.

Teacher Daily Behavior Card Data

As part of the universal intervention, the four teachers used a colored card system to track student behavior over the course of the day and to promote self-monitoring of behavior in their students. The system consists of five levels, each represented by a different color. The colors are ordered green, yellow, purple,

blue and red with green representing the most favorable behavior and red representing the most maladaptive behavior for a given day. The children were required to turn their cards to consecutive colors upon teachers' judgments of violation of classroom rules which include: 1) follow directions quickly; 2) raise your hand for permission to speak; 3) raise your hand for permission to leave your seat; 4) make smart choices; and 5) be respectful and kind. Thus, the data measured infraction frequency. This data was coded into a numerical system for data analysis in which green became a value of one, yellow became a value of two, purple became a value of three, blue became a value of four, and red became a value of five. The daily scores were summed and averaged across days that the student was present during the fall semester resulting in one outcome score with a value between one and five. This prevented absences from producing a misleadingly low score for daily behavior problems.

BASC™- 2 Behavioral and Emotional Screening System (BESS)

The *Behavioral and Emotional Screening System* (BESS) was utilized to assess parents' and teachers' perceptions of students' behavioral and emotional strengths and weaknesses (Kamphaus & Reynolds, 2007). Both the teacher and parent forms were administered using the child level appropriate for students in first grade at the beginning of the school year. The BESS provides one outcome score reflecting internalizing problems, externalizing problems, school problems, with adaptive skill items as well. Based on the normative sample, split-half reliability estimates for children aged 5-9 was in the mid .90s for both the Teacher and Parent Forms. Test-retest reliability was $r = .91$ for the Teacher Form and $r =$

.86 for the parent form. Inter-rater reliability of the Teacher and Parent Forms were $r = .70$ and $r = .87$ respectively utilizing ratings from two different teachers and two parents (i.e., mother and father) (Kamphaus & Reynolds, 2007).

Office disciplinary referrals

The school administrator provided ODR data; however, the teachers were inconsistent in their use of office referrals, with one teacher not distributing any. Therefore, this analysis was cancelled due to an inability to draw conclusions based on students' receipt of ODRs. Furthermore, evidence for the lack of technical adequacy of ODRs has been provided by McIntosh, Frank, and Spaulding (2010).

Procedure

Participants were recruited by contacting administrators and teachers in a district in the Southwestern US and communicating the benefits and costs associated with participating in the study. An incentive consisting of a gift-card to a department or craft store for the teachers' use in purchasing materials for the class was offered. The monetary amount was determined by adding five dollars for every student returning a parent consent form and parent report form of the BASC-2 BESS. At the beginning of the school year a packet including 1) parent consent forms, 2) documents describing the purpose of the study as well as risks and benefits of participation, and 3) a parent form of the BASC-2 BESS was sent home with each child. Upon the return of the parent materials for each individual child, the teachers were instructed to complete the BASC-2 BESS teacher form. On a predetermined day, the researcher visited the classroom to administer the performance tasks to each child on an individual basis, and each child was absent

from the classroom for no more than 10 minutes. Finally, at the end of the semester the teachers returned daily behavior card data to the researcher

Data Analysis

An a priori empirical power analysis was conducted using G* Power 3 computer software (2007) to estimate the ideal sample size N required for the statistical analyses intended to address the two research questions. The sample size N was computed as a function of power level $1-\beta$, significance level α , and a moderate effect size as defined by Cohen (1992). A power level set at .80, significance level set at .05, based on standard practice in the field, and an effect size of .15 or greater necessitated at least 77 participants for statistical significance. The effect sizes present in the current study were larger than assumed when calculating the a priori power analysis, which counterbalanced the smaller sample size and afforded the opportunity for the sufficient power requisite for statistical significance.

Results

A multiple regression analysis was conducted to evaluate how well the performance tasks and brief behavior rating scales (BBRSs) predicted teachers' observation of daily behavior in the classroom. The predictors were standardized tasks including the NEPSY-2 Auditory Attention task, NEPSY-2 Statue task, BASC-2 BESS Parent Form, and BASC-2 BESS Teacher Form. The criterion variable was an average score for each student's daily behavior level. Thus, the regression equation was $\hat{Y}_{\text{Classroom Behavior}} = \beta_0 + \beta X_{\text{Auditory Attention}} + \beta X_{\text{Statue}} + \beta X_{\text{Bess Parent}} + \beta X_{\text{Bess Teacher}}$. The linear combination of standardized tasks was significantly

related to the students' daily behavior, $F(4, 51) = 8.38, p < .01$. The sample multiple correlation coefficient was .63, indicating that approximately 35% of the variance in daily behavior can be accounted for by the linear combination of standardized behavior measures.

Indices to indicate the relative strength of the individual predictors is presented in Table 1. As expected, the performance tasks, for which better performance leads to higher scaled scores, were negatively related to the daily behavior scores, for which higher numbers reflect higher levels of maladaptive behavior. The BBRSSs, for which higher scores reflect higher perceptions of problematic behavior, were positively related to the daily behavior scores. All four standardized measures were significantly correlated with the daily behavior average ($p < .05$). The BASC-2 BESS Teacher Form was the only measure with a statistically significant partial correlation. However, conclusions about the relative importance of the screening tools are difficult because they are correlated. The correlations among the four standardized behavior measures were also statistically significant.

Another multiple regression analysis was conducted to evaluate whether the addition of multiple screening tools predicted students' daily behavior over and above the Teacher BESS, which showed the strongest correlation with student's daily behavior overall. The addition of multiple predictors did not account for a significant proportion of the variance after controlling for the effects of just the Teacher BESS, R^2 change = .05, F change (3, 51) = 1.18, $p = .33$. Thus,

using multiple screening tools appears to have little predictive value over using just the one with the strongest correlation with students' daily behavior.

A post-hoc multiple regression analysis was conducted to assess the extent to which the BASC-2 BESS Teacher form accounts for additional variance in daily behavior over and above the Parent form. The BASC-2 BESS Teacher form accounted for a significant portion of additional variance, R^2 change = .14, F change (3, 51) = 11.92, p = .001. Thus, the BASC-2 BESS Teacher form accounts for approximately 30% more variance when combined with the BASC-2 BESS Parent form. Additional details of the hierarchical model can be found in Table 2.

Chapter 5

Discussion

School Wide Positive Behavior Support (SWPBS) is a systematic problem solving approach that addresses the need for developing adaptive behaviors in students. The benefits of such models include the prevention of behavior problems, opportunity to intervene with children experiencing behavior problems, and increases in academic support (Forness, 2000; Kalberg, Lane, & Menzies, 2010). The first step in systematic problem solving models, such as Response to Intervention (RtI), includes universal screening for the purpose of identifying students in need of further intervention. Thus far, many of the screening tools developed for use in SWPBS, and those with the most empirical support, take the form of brief behavior rating scales (Kamphaus, et al., 2010; Kalberg, Lane, & Menzies, 2010). Furthermore, research contributing to best practices related to SWPBS is important to the notion of merging it with social, emotional, and behavioral forms of RtI. The current project investigated various psychometric tools for measuring children's behavior in an endeavor to identify the best instrument for use in screening children for additional behavioral intervention within a school utilizing SWPBS. First, a small study was conducted to obtain evidence that measures of children's behavior which involve the student performing some task requiring behavioral regulation, among other executive functions, are related to observers' ratings of children's externalizing behavior. Next, four instruments from two broad forms of assessment, rating scales and

performance tasks, were assessed for their validity and utility in SWPBS screening. The practicality of each instrument is also discussed.

Part I

The purpose of Part I of the study was to help determine which instrument(s) to use in Part II of the study. It is important to note that although the studied subtests have titles suggesting assessment of discrete executive functions, there may be construct irrelevant demands within the task and they therefore measure other related executive functions which may be difficult to discern based on performance. For example, there are aspects of inhibition in the Auditory Attention subtest, and aspects of attention in the Inhibition subtest.

Hypothesis 1

The expectation was that some tasks would be more highly related to reports of children's behavior, with reduced functioning on performance tasks associated with higher risk of behavioral problems. A relationship between a performance task and behavior ratings warranted consideration of the performance task for inclusion in research question two.

Results indicated that the Auditory Attention subtest of the NEPSY-II was significantly correlated with the Externalizing scale of the BASC-2, but the Inhibition subtest was not. This finding was similar to findings from other studies in which Lind, Haataja, Rautava, Väliäho, and Lehtonen (2010) found a significant correlation between the Auditory Attention subtest and parents' behavior ratings, whereas Hirayama (2011) did not find a significant relationship between the Inhibition subtest and parent rating scales. One explanation for the

Inhibition subtest not predicting externalizing behavior as one would expect could be the construct irrelevant skills required to perform well, such as rapid automatic naming. It is noteworthy that in both of the previously mentioned studies subtest scores were reduced in children with emotional or behavior disorders; however, the NEPSY-II Technical manual reports that children with such problems showed lower performance than matched controls from the standardization sample across several domains of functioning including those seemingly unrelated to externalizing behavior (Kemp & Korkman, 2007).

Part II

The second part of the study involved implementation of various screening tools within an elementary school in the beginning stages of implementing a SWPBS program. The tools used came from two broad categories: brief behavior rating scales and performance tasks.

Screening tools

There has been an apparent increase in the production of BBRs recently, presumably to meet the needs of efficiency and psychometric adequacy of screening tools in SWPBS. A challenge associated with BBRs as screeners include being sensitive enough to detect behavior problems before they arise in the case of prevention, and without delay in the case of intervention. Additionally, they must be brief enough to be practical for teachers to use (Kamphaus, et al., 2010). As with most forms of assessment, psychometric properties such as practicality and validity of SWPBS screeners should be maximized to make them more useful to educators. This means that they need to be as short as possible,

while also being relatively reliable and ecologically valid. Teachers and other education staff face many demands on their time, making it essential that screening for academic and behavioral risk factors be as efficient as possible. The current study assessed two versions of the Behavioral Assessment Scales for Children, Second Edition Behavioral and Emotional Screening System (BASC-2 BESS). Although BBRs are the most predominant tools available for screening in SWPBS, they are not the only choice. The performance tasks that have been developed as direct measures of skills related to children's behavioral regulation also seem useful.

Hypothesis 2

It was expected that teachers' rating scale responses would have a more modest relationship with both their later reports and with office disciplinary referrals than the parent reports and performance tasks. This hypothesis was not substantiated, in part because it was based on an assumption that all tools would be administered simultaneously. However, standard administration requires that teachers know the students for at least one month before rating them. In the current study, the parent BESS forms and teacher BESS forms were administered approximately three to four weeks apart. It was assumed that the three to four week delay in administration would not be a significantly confounding variable due to the high test retest reliability established in the instruments standardization (Kamphaus & Reynolds, 2007). However, this delay may be practically significant since parental ratings were conducted during the summertime, prior to

enrollment for some students in any full time school program and teacher ratings were conducted during the first month of the school year.

Independent measures

BASC-2 BESS Teacher Form.

In the current study, the BASC-2 BESS completed by teachers showed the strongest correlation with their ratings of children's behavior on a daily basis over an entire semester with a large effect size (Cohen, 1993). This piece of evidence might suggest that it is therefore the best single instrument to use when screening for behavior problems in young children in order to provide prevention services or early intervention. Burke et al. (2012) also found the BASC-2 BESS to be strongly related to teacher's expectations within SWPBS. However, certain aspects of the instrument and its relationship with the outcome measure make its stance as the best tool for use in SWPBS arguable. When choosing a screening instrument, it is important to consider practicality in addition to predictive validity, whereas the BASC-2 BESS teacher form requires more personnel time than any of the other of the instruments investigated in the current study.

Standardized administration of the BASC-2 BESS requires that the teachers know the student for at least 4 weeks before completing the rating scale. This was the longest delay after the beginning of the school year of the instruments used in this study, and therefore would be the least conducive to prevention or early intervention. Yet, one of the advantages of three tiered models of service delivery is the opportunity to prevent problems from arising in the first place. Additionally, because the measure relies on teachers' perceptions of problem behaviors, it is not

ideal for prevention efforts, which Burk et al. (2012) state is an important objective in SWPBS, because presumably problematic behaviors would already have begun to influence teachers' perceptions. Furthermore, because both measures rely on the same individual's perception of the construct (student behavior), it is possible that a form of behavioral confirmation (Gross, 1983) has influenced the high correlation between the two which would be a confounding variable. Finally, before investing in this instrument it would be useful to know the extent to which it performs better at predicting children's need for intervention over and above simple teacher nomination. The BASC-2 BESS Teacher Form was highly correlated with teachers' daily rating of children's behavior; however, it is easy to identify aspects of the instrument which make its efficiency questionable.

BASC-2 BESS Parent Form.

The BASC-2 BESS Parent Form also had a convincing correlation with teachers' daily ratings of student behavior, although not to the extent of the Teacher Form, with a medium to large effect size (Cohen, 1993). An important aspect of the measure is that it does not require teachers' time to administer, although teachers in some settings will likely be responsible for their dissemination and collection. School administrators and teachers alike will probably appreciate that the instrument saves personnel resources. One aspect of the BASC-2 BESS Parent Form which may be problematic in authentic universal screening for SWPBS, and was not able to be addressed in the current study because of issues related to confidentiality of data, is the notion that parents may

be reluctant to share information about their children's problematic behavior up front for fear that it will have a negative impact on their child. The Institutional Review Board responsible for the current study required notification about participants' confidentiality in the consent forms, which likely interfered with any conclusions one could draw about parents' hesitation to share their perception about their child's problem behavior because parents were told that the results of their forms would not be shared with anyone apart from the researcher. However, the earlier knowledge of maladaptive behaviors afforded by the parent version of the BESS seems valuable as it allows for earlier intervention and possibly prevention of problematic school behaviors.

NEPSY-II Statue Subtest.

The correlation between the NEPSY-II Statue subtest and the daily behavior ratings also had a medium to large effect size (Cohen, 1993), which was almost identical to that of the BASC-2 BESS Parent Form, which makes it another significant predictor of students needing behavioral intervention. The benefits of this instrument is that it can be given immediately within a school year, takes very little time to administer and score (less than three minutes), and does not rely on other individual's perceptions. The downside of the instrument is that it requires a certain amount of training to administer, although to a school psychologist this training would seem negligible.

NEPSY-II Auditory Attention.

The NEPSY-II Auditory Attention Subtest had a small to medium sized correlation with the daily behavior ratings, which was the weakest of the four

instruments investigated. Additionally, it takes a moderate amount of time to administer and requires a relatively higher level of training for an examiner. Thus, it shows an unimpressive level of predictive validity and does not seem to be very practical compared to the other instruments used in the study.

Office disciplinary referrals (ODRs)

In the current study the teachers' use of office disciplinary referrals was sparse and inconsistent to such a degree that they could not be used in statistical analysis. Unfortunately for researchers, this is not an uncommon pattern in schools (McIntosh, Frank, & Spaulding, 2010). Furthermore, even studies in which ODR data was collected in a standardized fashion, low levels of reliability and relation with SWPBS expectations were found (Burke et al., 2012).

Hypothesis 3

It was expected that the addition of multiple tools upon universal screening would enhance the ability to predict behavioral problems as determined at the end of the semester. This hypothesis was substantiated, but only minimally. When additional subtests were added to the Teacher BESS, which was the measure most strongly correlated with the daily behavior outcome measure when considered alone, it did not account for a statistically significant increase in variance. Thus, the addition of multiple measures does not seem efficient in SWPBS screening procedures.

Due to the high correlation between the Parent BESS and Teacher BESS, a post hoc analysis was conducted to investigate the increase in variance accounted for when the Teacher BESS was added to the Parent BESS because it is

believed that the Parent BESS is a more efficient measure to use in screening because 1) it does not require teacher time to administer and 2) it can be administered earlier in the school year. The BASC-2 BESS Teacher form accounted for a statistically significant portion of additional variance, approximately one third more. Thus, it is not clear if the parent rating scale is adequate for prediction of student classroom behavior. At this time, if choosing one rating scale to administer, the BASC-2 BESS Teacher form seems the most useful. However, because the Teacher BESS and the daily behavior outcome measure used in these analyses were both conducted by teachers, it is possible that a Pygmalion effect is influencing the outcome through common variance (Rosenthal & Jacobson, 1992).

Limitations

One of the limitations associated with Part I of the study includes a lack of data for a wider selection of subtests from the Executive Functioning domain of the NEPSY-II. Furthermore, the archival nature of the data interferes with the assumption of random sampling requisite in multiple regression analysis. Part II of the study was conducted with limitations related to a sample restricted to a specific geographic location and school climate. Therefore, results may not generalize to populations across the United States or to schools that do not hold a Title I designation. Furthermore, aspects of the administration of the BASC-2 Parent Form were not authentic to a real-world situation given that parents were informed that their child's data would be kept confidential and not shared with anyone. This would not be the case at a school that was actually using the

instrument as a screening tool, and therefore conclusions about its utility as a universal screener in SWPBS should be interpreted with caution. Furthermore, comparisons made between various instruments' abilities to predict the daily behavior outcome measure should be considered unclear because of the connection between the Teacher BESS and daily behavior score in that they were both rated by the teacher and are therefore expected to have a relationship moderated by teachers' expectations.

Future Directions

Future research on this topic should include addressing the current study's limitations. More specifically, the study should be conducted again with students from other parts of the United State and of various socio-economic statuses. The weaknesses of the individual instruments should also be investigated further. For example, it would be interesting to know if parents' willingness to share their perceptions of their children's behavior would change if they believed the information would be shared with his or her teacher. If parents are unwilling to share information about their child, the utility of the BASC-2 BESS would be diminished. However, it is possible that other school personnel, such as a school psychologist, could manage this data making it unnecessary for parents' perceptions of their children to be reported to their teachers.

Regarding the BASC-2 BESS Teacher Form, it would be useful to know if the extra time required to fill out the form leads to greater predictive validity beyond what could be achieved with simple teacher nomination. Studies utilizing outcome data that does not rely on teacher perceptions would also be useful.

Finally, there are many other instruments including BBRS's and performance based tasks that were not included in the current study, but which may show some utility as universal screeners within SWPBS. Examples of such instruments from the BBRS classification are the BIMAS (McDougal, Bardos, & Meiers, 2011) and DESSA-Mini (LeBouf, Nagleiri, & Shapiro, 2010), especially because they are designed to be used as progress monitoring tools as well. The assessment of those instruments' psychometric properties may be helpful to educators. Alternatively, it may be appropriate to develop performance tasks specifically for the purpose of screening rather than borrowing psychometric tools created for other purposes such as comprehensive psychological and neuropsychological evaluations.

Conclusion

Universal screening procedures can contribute to improved behavioral health in schools (Albers, Glove, & Kratochwill, 2007). This is accomplished through early identification of students needing additional behavioral intervention, and in most cases the earlier that identification can be achieved, the better the outcome (Forness et al., 2000). However, schools generally have limited resources such as money and personnel time. Therefore, it is important that an individual screening tool be as practical as possible, with respect to how soon it can be administered and what resources are necessary for its use, while maintaining the integrity of its predictive validity. Within the current study, three possible screening procedures showed promise, but with varying levels of practicality and predictive abilities. Future research initiatives may help to resolve concerns regarding the individual instruments, but in the interim school systems

should choose screening procedures most conducive to their own resources to prevent the burden of universal screening from interfering with the fidelity of three tiered problem solving models.

REFERENCES

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-aged forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth and Families.
- Albers, C. A., Glover, T. A., & Kratochwill, T. R. (2007). Where are we, and where do we go now? Universal screening for enhanced educational and mental health outcomes. *Journal of School Psychology, 45*, 257-263.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Briesch, A. M., Chafouleas, S. M., & Riley-Tillman, T. C. (2010). Generalizability and dependability of behavior assessment methods to estimate academic engagement: A comparison of systematic direct observation and direct behavior rating. *School Psychology Review, 39*, (3), 408-421.
- Brooks, B. L., Sherman, E. M. S., & Strauss, E. (2010). Test review: NEPSY-II: A developmental neuropsychological assessment, second edition. *Child Neuropsychology, 16*, 80-101.
- Burke, M. D., Davis, J. L., Lee, Y., Hagan-Burke, Shanna, Kwok, O., & Sugai, G. (2012). Universal screening for behavioral risk in elementary schools using SWPBS expectations. *Journal of Emotional and Behavioral Disorders, 20*(1), 38-54. doi: 10.1177/1063426610377328
- Cohen, J. (1992). A power primer, *Quantitative Methods in Psychology, 12*, 155-159.
- Davidson, R. J. (2000). Affective style, psychopathology, and resilience: Brain mechanisms and plasticity. *American Psychologist, 55*, 1196-1214.
- Davis, J. L. & Matthews, R. N. (2010). NEPSY-II review. *Journal of Psychoeducational Assessment 28* (2), 175-182.
- Derryberry, D. & Rothbart, M. K. (1988). Arousal, affect, and attention as components of temperament. *Journal of Personality Social Psychology, 55*, 958-966.
- Dishion, T. J., & Patterson, G. R. (1999). Model building in developmental psychopathology: A pragmatic approach to understanding and intervention. *Journal of Clinical Child Psychology, 48* (4), 5012-512.

- Eisenberg, N., Fabes, R. A., Guthrie, I. K., Murphy, B. C., Maszk, P., Holmgren, R., & Suh, K. (1996). The relations of regulation and emotionality to problem behavior in elementary school children. *Development and Psychopathology, 8*, 141-162.
- Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Lesoya, S., Murphy, B. C., Jones, S., Poulin, R., & Reiser, M. (2000). Prediction of elementary school children's externalizing problem behaviors from attentional and behavioral regulation and negative emotionality. *Child Development, 71*(5), 1367-1382.
- Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.
- Fellers, C. L., Duhon, G., Giblet, M.A., Mocco, C., Knepper, K., Villanueva, C. S., & Fontenelle, S. (February 2011) *Response to Intervention for Behavioral Concerns*. Poster presentation at National Association of School Psychologists Conference in San Francisco, CA.
- Forness, S. R., Serna, L.A., Nielson, E., Lambros, K., Hale, M.J., & Kavale, K.A. (2000). A model for early detection and primary prevention of emotional and behavioral disorders. *Education & Treatment of Children, 23*, 325-346.
- Gallo, F. J. (2009). *Executive Functions in Young Children with Williams Syndrome*. ProQuest Information & Learning. AAI3363432.
- Garcia-Berrera, M. A., Kamphaus, R. W., & Bendalos, D. (2011). Theoretical and statistical derivation of a screener for the behavioral assessment of executive functions in children. *Psychological Assessment, 23* (1), 64-79.
- Gresham, F. M., Cook, C. R., Collins, T., Dart, E., Rasetshwane, K., Treulson, E., & Grant, S. (2010). Developing a change-sensitive brief behavior rating scale as a progress monitoring tool for social behavior: An example using the Social Skills Rating System- Teacher Form. *School Psychology Review, 39* (3), 364-379.
- Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

- Hale, J. B., & Fiorello, C. A. (2004). *School Neuropsychology: A Practitioner's Handbook*. New York: The Guilford Press.
- Hart, D., Keller, M., Edelstein, W., & Hoffman, V. (1998). Childhood personality influences on social-cognitive development: A longitudinal study. *Journal of Personality and Social Psychology*, *74*, 1278-1289.
- Hinshaw, S.P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, *111*, 127-155.
- Hirayama, K. L. (2011). *The Relationship between Inhibition and Externalizing Behavior in Boys with Severe Behavior Problems*. ProQuest Information & Learning. AAI3447078.
- Individuals with Disabilities Education Improvement Act of 2004*, 20 U.S.C. 1400 et seq. (2004). (reauthorization of Individuals with Disabilities Act 1990).
- Kamphaus, R.W., DiStefano, C., Dowdy, E., Eklund, K., & Dunn, A. Determining the presence of a “problem”: Comparing two approaches for detecting youth behavioral risk. *School Psychology Review*. 2010. [Refereed Journal Article]
- Kamphaus, R. W., & Reynolds, C. R. (2007). *Behavioral and Emotional Screening System*. Bloomington, MN: Pearson Clinical Assessment.
- Kalberg, J. R., Lane, K. L., & Menzies, H. M. (2010). Using systematic screening procedures to identify students who are nonresponsive to primary prevention efforts: Integrating academic and behavioral Measures. *Education and Treatment of Children*, *33* (4), 561-584.
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY- Second Edition (NEPSY-II)*. San Antonio, TX: Harcourt Assessment.
- Lane, K. L., Gresham, F.M., & O’Shaughnessy, T. E. (2002). Serving students with or at-risk for emotional and behavioral disorders: Future challenges. *Education and Treatment of Children*, *25* (4). 507-521.
- Naglieri, J., LeBuffe, P., & Shapiro, V. (2010). *Devereux Student Strengths Assessment –Mini*. Lewisville, NC: Kaplan
- Lind, A., Haataja, L., Rautava, L., Väliäho, A. Lehtonen, Li. (2010) [Relations between brain volumes, neuropsychological assessment and parental questionnaire in prematurely born children](#). *European Child & Adolescent Psychiatry*, *19* (5), 407-417.

- McDougal, J. L., Bardos, A. N., & Meier, S. T. (2011). *Behavior Intervention Monitoring System*. Multi-Health Systems: ON, Canada.
- McIntosh, K., Campbell, A. L., Carter, D. R., & Zumbo, B. D. (2009). Concurrent validity of office discipline referrals and cut points used in Schoolwide Positive Behavior Support. *Behavior Disorders, 34*, 100-113.
- McIntosh, K., Frank, J. L., & Spaulding, S. A. (2010). Establish research-based trajectories for office discipline Referral for individual students. *School Psychology Review, 39*, 364-379.
- National Institute of Child Health and Human Development. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). *Report of the National Reading Panel*. Washington, D.C: US Government Printing Office.
- Pas, E. T., Bradshaw, C. P., & Mitchell, M. M. (2011). Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychology in the Schools, 48*(6), 541-555.
- Patterson, G. R., & Newman, J. P. (1993). Reflectivity and learning from aversive events: Toward a psychological mechanism for the syndromes of disinhibition. *Psychological Review, 100*, 716-736.
- Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47*, 427-469.
- Reynolds, C. R. & Kamphaus, R. W. (2004). *Behavior Assessment System for Children (2nd Ed.)*. Circle Pines, MN: American Guidance Service.
- Riccio, C. A., Sullivan, J. R., & Cohen, M. J. (2010). *Neuropsychological Assessment and Intervention for Childhood and Adolescent Disorders*. New Jersey: John Wiley & Sons.
- Rosenthal, R. & Jacobson, L. (1992). *Pygmalion in the classroom* (Expanded ed.). New York: Irvington.
- Sadler, C., & Sugai, G. (2009). Effective behavior and instructional support. *Journal of Positive Behavior Interventions 11* (1). Pages.
- Saeki, E., Singh, R., Renshaw, T., Earhart, J., Stewart, K., Shane, J., & Hart, S. (February 2011) *RtI in the Social, Emotional, and Behavioral Domains: Current Challenges and Emerging Possibilities*. Paper Presentation at

National Association of School Psychologists Conference in San Francisco, CA.

- Skinner, B. F. (1953). *Science and human behavior*. New York: MacMillan.
- Satcher, D. (2001). *Youth Violence: A Report of the Surgeon General*. Washington, DC: U.S. Public Health Services, U.S. Department of Health and Human Services.
- Sugai, G., & Horner, R. H. (2008). What we know and need to know about preventing problem behavior in schools. *Exceptionality, 16*, 67-77.
- Tan, C. S. (2007). Test review: Reynolds, C. R. & Kamphaus, R. W. (2004). Behavior assessment System for children (2nd ed.). Circle Pines, MN: American Guidance Service. *Assessment for Effective Intervention, 32* (2).
- Tobin, T.J., & Sugai, G.M. (1999). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional & Behavioral Disorders, 7*, 41-55.
- Tobin, T.J., Sugai, G.M., & Colvin, G. (1996). Patterns in middle school discipline records. *Journal of Emotional and Behavioral Disorders, 4*, 82-94.
- U. S. Department of Education, Office of Special Education Services. (2010). *Implementation Blue-Print and Self Assessment: School-Wide Positive Behavioral Supports and Interventions*. Technical Assistance Center on Positive Behavioral Interventions and Supports. Retrieved from <http://www.cde.state.co.us/pbis/download/pdf/Blueprint.pdf> on June, 2nd 2011.
- Vollmer, T. R., & Iwata, B. A. (1992). Differential reinforcement as treatment for behavioral disorders: Procedural and functional variations. *Research in Developmental Disabilities, 13*, 393-417.
- Volpe, R. J. & Gadow, K.D. (2010). Creating abbreviated rating scales to monitor classroom inattention-overactivity, aggression, and peer conflict: Reliability, validity, and treatment sensitivity. *School Psychology Review, 39*, 350-363.
- Walker, B. A. (2010). Effective schoolwide screening to identify students at risk for social and behavioral problems. *Intervention in School and Clinic, 46* (2), 104-110.

APPENDIX A

TABLES

Table 1
Correlations of Screening Instruments and Daily Behavior Outcome

	Parent BESS	Teacher BESS	Statue	Auditory Attention
Daily Bx Average	.42**	.60**	-.43**	-.28*
Parent BESS		.52**	-.34*	-.32*
Teacher BESS			-.46**	-.37*
Statue				.38*

* $p < .05$


** $p \leq .001$

Table 2
Hypothesis 3 Hierarchical Regression

Predictor	ΔR^2	β
Step 1		
Parent BESS	.19*	.67*
Step 2		
Teacher BESS	.14*	1.08*
Total R^2	.57*	

* $p \leq .01$

APPENDIX B
HUMAN SUBJECTS IRB APPROVAL



Office of Research Integrity and Assurance

To: Linda Caterino Kulhavy
EDB

MR
From: Mark Roosa, Chair
Soc Beh IRB

Date: 07/20/2011

Committee Action: Expedited Approval

Approval Date: 07/20/2011

Review Type: Expedited F7

IRB Protocol #: 1106006605

Study Title: Validity and Practicality of Screening in SWPBS: Comparing Three Methods

Expiration Date: 07/19/2012

The above-referenced protocol was approved following expedited review by the Institutional Review Board.

It is the Principal Investigator's responsibility to obtain review and continued approval before the expiration date. You may not continue any research activity beyond the expiration date without approval by the Institutional Review Board.

Adverse Reactions: If any untoward incidents or severe reactions should develop as a result of this study, you are required to notify the Soc Beh IRB immediately. If necessary a member of the IRB will be assigned to look into the matter. If the problem is serious, approval may be withdrawn pending IRB review.

Amendments: If you wish to change any aspect of this study, such as the procedures, the consent forms, or the investigators, please communicate your requested changes to the Soc Beh IRB. The new procedure is not to be initiated until the IRB approval has been given.

Please retain a copy of this letter with your approved protocol.