Theses, Dissertations and Capstones

1-1-2013

Concurrent and Predictive Validity of the Behavior and Emotional Screening System

Amber Rae Wallbrown
amberwallbrown@yahoo.com

Follow this and additional works at: http://mds.marshall.edu/etd

Part of the School Psychology Commons

CONCURRENT AND PREDICTIVE VALIDITY OF THE BEHAVIOR AND EMOTIONAL
SCREENING SYSTEM

A thesis submitted to
the Graduate College of
Marshall University

In partial fulfillment of
the requirements for the degree of
Education Specialist

in

School Psychology

by
Amber Rae Wallbrown

Approved by
Dr. Sandra S. Stroebel, Committee Chairperson
Dr. Fred Jay Krieg
Dr. Stephen L. O'Keefe

Marshall University
August 2013

CONTENTS

LIST OF TABLES

Abstract

Behavior Screeners are important tools for early identification of children's social, emotional, and behavioral problems in schools. This study evaluated the concurrent and predictive validity of the BASC-2 Behavioral and Emotional Screening System (BASC-2 BESS). It was compared to the Behavior Assessment System for Children, Second Edition (BASC-2) to determine the congruent validity between the instruments. Predictive validity was examined by comparing the instruments to office referrals. The rating scales were administered to parents (96% mother, 4% father) of 8 identified and 15 non-identified students (mean age of 10.3 years; 52% male, 48% female). Pearson correlation coefficients examining the consistency between the two instruments were generally strong and positive, supporting the use of both instruments for assessing a child for behavioral or emotional problems. An analysis of hits and misses found the BASC-2 BESS accurate for predicting office referrals and therefore usable for screening students for possible behavioral problems.

Chapter 1

Literature Review

Student behavior in school settings receives substantial attention within practice and research arenas due to recent high profile acts of violence and antisocial behavior in the school environment (Chafouleas, Kilgus, Maggin, & Sanetti, 2012). Identifying youth at-risk prior to the development of, or during the initial stages of problem behaviors, is necessary for effective implementation of early intervention designed to prevent negative adolescent outcomes such as conduct problems, substance use, and poor school adjustment (Boelter, Caldarella, Gentry, Merrell, & Streeter, 2001). When behavior problems are suspected, rating scales and screeners can help with identifying the behavior problem and subsequently with monitoring of improvement.

Early Detection of Behavior Problems

Young children who lack social and emotional competence demonstrate discipline problems frequently in school and are at risk for not achieving future academic and professional success (Barber, Richardson, Schultz, & Wilcox, 2011). Research has indicated that deficits in social skills are often predictive of juvenile delinquency and referral for mental health services. Research has found that childhood aggression and peer rejection are significant predictors of psychological disorders during adolescence. One of the primary causes of poor peer relations experienced by youth who exhibit conduct problems is aggressive behavior (Boelter et al., 2001). Conduct problems increase the risk of substance abuse, delinquency, violent behavior, and school dropout (Hawkins, Catalano, Kosterman, Abbott, & Hill, 1999). Young children who become aggressive may establish a pattern of social problems by the time they reach lower

elementary school. Many children who demonstrate behavioral problems are being dismissed from schools by the time they are six years old. In order for interventions to reduce or prevent behavioral problems, they must begin at an early age (Barber et al., 2011). Identifying problem behaviors occurring in the home, school, and community during childhood is crucial in the prevention of negative adolescent and adult social outcomes (Boelter et al., 2001).

Prevention of Behavior Problems in the School Setting

Research has documented the important role school settings serve in the prevention of behavior problems. Behavior screeners are frequently used in school and community settings as a means of early identification of children's social, emotional, and behavioral problems. Comprehensive Behavior rating scales are used for children referred or identified as having problem behaviors. If there is not early identification of the problem behaviors and implementation of preventative interventions, aggressive, antisocial behaviors demonstrated by children may be predictive of severe lifelong problems (Boelter et al., 2001).

One important component of prevention and early intervention is early detection. Screening children helps determine those who are having difficulties or are at risk for difficulties so problems can be addressed early on. Early intervention is needed to make the most impact. Another important component is continual evaluation of behavioral change in response to implemented supports. This continual process requires behavior assessment tools that are both technically adequate for their intended purposes and quick and easy to use for their intended users (Chafouleas et al., 2012).

Instruments to Evaluate Behavior and Emotional Functioning in Youth

  Behavior Screeners. Comprehensive behavior rating scales are an option for discovering

the presence of behavioral and emotional problems, but they are not well suited for large-scale

testing due to the time and effort that is required to assess hundreds or thousands of children

within a given school district. Behavior Screeners offer a quick and inexpensive way to identify

children with possible problems. A significant barrier to the establishment of behavioral and

emotional screening programs in schools, hospitals, and pediatric practices is due to the lack of

associated evidence of reliability and validity with these practical screening tools. The BASC-2

Behavioral and Emotional Screening System (BASC-2 BESS) was developed to fill the void in

behavioral and emotional screening in children. The BASC-2 BESS offers schools and mental

health professionals a reliable, quick, multifaceted and economical way for early detection of

students who may be in need of additional services because large numbers of children can be

administered the screening to determine if there are any potential problems (Kamphaus &

Reynolds, 2007).

  Implementing screening systems for behavioral and emotional problems provides schools

with an objective, efficient, and systematic way to identify children that may develop academic

or other school-related problems. Introducing a systematic screening approach will benefit a

school district by allowing: (a) quick evaluation of all children within a grade or school, instead

of only those children referred by a teacher or other school staff; (b) early identification,

allowing schools to catch problems before they become too serious; and (c) a standardized way

of identifying students that have a high likelihood of experiencing school-related problems

because of behavioral or emotional difficulties (Kamphaus & Reynolds, 2007).

The terms "screening" and "assessment" are not interchangeable. Screening is a preliminary process that identifies, from all the children, those who may be at risk for future difficulty in school and those who may have special needs in learning. In both cases, those identified children must be assessed more carefully to evaluate whether they do require adaptations of the regular instructional program, or qualify for specialized educational placement (Kamphaus & Reynolds, 2007). Screening provides psychologists with information on which children need a more time consuming and expensive evaluation. Screeners indicating problems will often be followed up with behavior rating scales.

Behavior Rating Scales. Behavior rating scales are useful for more in depth evaluations of children. Information gathered from behavior rating scales is used for several purposes including screening, diagnosis, or monitoring the effects of treatment (Angello, DiPerna, Gureasko-Moore, Gureasko-Moore, Nebrig, Ota, & Volpe, 2003). They are also used to plan behavioral interventions and support plans and to monitor the effectiveness of the interventions or support plans over time (Hosp, Hosp, & Howell, 2003). Best practice in behavioral assessment requires practitioners to gather information using a variety of assessment methods completed by multiple sources describing behaviors across different settings. Because behavior rating scales provide a time and cost effective way to obtain parent and teacher perceptions of the presence and severity of a child's behaviors in a broad range of problem areas, school psychologists include these rating scales as part of multisource, multimethod assessments (Bergeron, Farmer, Floyd, McCormack, 2008). Behavior rating scales typically list dozens of behavioral descriptors. The informant marks their perceptions as to the degree to which each behavior is present in the child (Bour, Hakman, Murphy, Myers, & Sidebottom, 2010). So although they are not overly time consuming, the 30 minutes required to rate a student does not make the wide scale usage for

all children in a school or hospital setting practical. Rating scales are often reserved for children who are referred or identified as having problems.

One rating scale commonly used in schools is the Behavior Assessment System for Children Second Edition (BASC-2). It is a standardized multi-dimensional rating system that assesses a broad range of skills, adaptive behaviors, and problem behaviors using child self-reports, teacher ratings, and/or parent ratings. The BASC-2 items are oriented toward DSM-IV symptomology, and can be used in screening and as a comprehensive assessment (Knoll et al., 2010). Behavior rating scales (a) provide quantifiable information, which can be held to standards of reliability and validity; (b) provide systematically organized information; (c) are efficient to complete and score; (d) include normative data, which allow comparison of individual behaviors to that of large groups; and (e) can be used to compare ratings of different respondents or across settings (Hosp et al., 2003). The BASC-2 has been described as one of the best behavior-rating scales (Bour et al., 2010).

Parents as Raters

When conducting screenings and assessments, school-based clinicians seek information frequently from parents/caregivers using rating scales. Parents are considered a critical source of information related to cross-setting social and behavioral performance because of the comprehensive information they possess (Knoll, Lopata, Rodgers, Smerbeck, Thomeer, Toomey, & Volker, 2010). Parents are typically among the most important sources of data about children's competencies and problems. Noted in a recent survey of psychological assessments of adolescents by psychologists in general, one change in test usage has been that of an increased use of parent rating scales (Blaha, Merydith, & Prout, 2003). In a study examining the clinical

and adaptive features of students with high-functioning autism spectrum disorder (HFASD), evidence suggested that the BASC-2 PRS detected important differences between children with HFASD and typically-developing children on scales assessing both core and associated features of ASD (Knoll et al., 2010). Some mental health diagnoses require a child to exhibit problems in more than one setting. Thus, there is the need to obtain information on how children are functioning in the home.

Considerations when Selecting an Appropriate Instrument

There are a variety of rating scales and screeners available for use by education professionals, clinicians, and researchers, ranging from those designed to assess social skills to those designed to assess specific disorder-related problem behaviors. Because of the need for data to be gathered from multiple sources and across multiple settings, it is imperative that researchers develop comprehensive assessment tools that can be used in such ways (Boelter et al., 2001). Given the variety of ratings scales currently available, selecting an appropriate rating scale or screener can be a confusing task (Angello et al., 2003). Several considerations should be taken into account when deciding if a specific behavior rating scale or screener is appropriate for a certain purpose (Hosp et al., 2003). The psychometric integrity of instruments and the adequacy of normative data with respect to the child's culture in addition to the purpose of assessment and age of the child are essential considerations when selecting an appropriate instrument (Angello et al., 2003).

Commercially available rating scales and screeners are generally well validated and exhibit exceptional psychometric characteristics (Gadow & Volpe, 2010). A key element to an instrument's reliability and validity is the consistency of ratings (Bour et al., 2010). Validity

determines the degree of accuracy a test has for a specific purpose. Reliability determines the consistency, or stability, in measurement (Hosp et al., 2003). To determine construct validity, test publishers include results in the tests' manuals as to how specific scales correlated with scales on other instruments (Bour et al., 2010). High correlations between instruments designed to measure the same or like constructs render convergent evidence. Low or non-significant correlations among measures lacking a conceptual relationship render divergent evidence (Blaha et al., 2003). Even though publishers need to examine and report on the validity and reliability of their instruments, users of these instruments need to evaluate them for themselves. Although research has been done on the behavior rating scales, information on the newer behavior screeners is more limited.

Current Research

Behavior Assessment System for Children Second Edition (BASC-2). The BASC-2 has been compared to several related behavioral assessment tools: the Achenbach System of Empirically Based Assessment Caregiver-Teacher Report Form, the Conners' Teacher Rating Scale-Revised, and the previous version of the BASC. Correlations between subscales were high (in the .70 and .80s) when they addressed similar content (Barber et al., 2011). Validity evidence for the scores obtained from the Social Skills Improvement System—Rating Scale has been demonstrated by correlational studies with the BASC-2 (Cook, Elliot, Gresham, Kettler, & Vance, 2010). In a study designed to compare the consistency of two behavior rating scales, all 15 comparisons of interest between corresponding scales on the BASC-2 PRS and Child Behavior Checklist (CBCL) were positive and significantly correlated (Bour et al., 2010). Another study found the pattern of correlations between the Home and Community Social Behavior Scales (HCSBS) and the BASC-PRS remarkably similar across the two age range

forms. Correlations between the HCSBS Social Competence Scale and the similar subscales from the BASC were moderate to strong, ranging from .51 to .78 across the two age range forms (Boelter et al., 2001). In a study examining the behavioral profiles of children with epilepsy, the results demonstrated good agreement with previous research using the Achenbach CBCL and TRF (Teacher Report Form), offering support for the use of the BASC-2 PRS and TRS for assessing behavioral functioning in the children with epilepsy (Blackburn, Sanders, Kanive, & Titus, 2008). A study evaluating the relative sensitivity to change, the parent versions of the CBCL, BASC-2, and Youth Outcome Questionnaire (Y-OQ) were positively correlated and are recommended for use over the self-report versions of the measures when data from only one informant can be obtained (Burlingame, Eggett, McClendon, McClendon, Warren, & Green, 2011).

BASC-2 Behavioral and Emotional Screening System (BASC-2 BESS). A search for reliability and validity studies on the BASC-2 BESS yielded little results. The BASC-2 BESS proved to be a diagnostically accurate predictor of BASC-2 Externalizing Problems and School Problems risk. BASC-2 BESS predictive validity was supported through a 2-year longitudinal study of K-5 students in which the BASC-2 BESS was associated with a variety of student outcomes including conduct problems, social skills, special education placement, and grades (Chafouleas, Kilgus, Riley-Tillman, & Welsh, 2012). Information from the manual indicates Total Scores from the Teacher, Parent, and Student Forms were compared with the results obtained from the following instruments used in the correlational studies: Behavior Assessment System for Children, Second Edition, Achenbach System of Empirically Based Assessment, Conners' Rating Scales-Revised, Vineland Adaptive Behavior Scales, Second Edition, Behavior Rating Inventory of Executive Function, Children's Depression Inventory, and Revised

Children's Manifest Anxiety Scale. The relationships between the Total Score and the BASC-2 global composite scores are very strong, with correlations ranging from .86 to .94. Strong relationships also exist between the Total Score and the other composite scores offered on the BASC-2, providing evidence that the Total score is predictive of a wide variety of behavioral and emotional problems. The correlations of the Total Score with the ASEBA Total Problems composite scales are strong across forms, ranging from the lower to upper .70s. This provides further evidence for the strong relationship between the Total Score and other measures of behavioral and emotional problems. The correlations show that the Conners' Global Index and the ADHD-specific indices have strong relationships with the Total Score, ranging from the lower .60s to the mid-to upper .70s, which provides evidence that the Total Score is a good indicator of behaviors associated with ADHD. Strong correlations were obtained between the Total Score and the global indices on the BRIEF, .78 and .64, which indicate that the Total Score is predictive of executive functioning problems. Correlations between the Total Score and the Vineland ABC were strong, -.66, to moderate, -.50, which suggest at least a moderate relationship between the Total Score and adaptive behavior. The correlations obtained between the Total Score and overall scores from the CDI and RCMAS were moderate, .50s. These indicate a modest predictor of depression and anxiety (Kamphaus & Reynolds, 2007).

Whereas research conducted by publishers is important and guides test development, there is the need for research in the field to validate instruments with different populations. A computerized search of PsychINFO and PsychARTICLES using Academic Search Premier based search using the terms Behavior Assessment System for Children Second Edition and BASC-2 Behavioral and Emotional Screening System, no studies were found comparing the two

instruments. The current study attempts to provide more information about the relationship between the BASC-2 BESS and the BASC-2.

Purpose of Current Study

The purpose of this study was to examine how consistently the parent version of a behavior rating scale and a behavior screener measured similar constructs in identified and non-identified students. Consistent measurement of those constructs would suggest that they measure the same types of behaviors (Bour et al., 2010) and help establish the validity of the BASC-2 BESS as a psychometrically sound and useful behavior screener (Boelter et al., 2001). The purpose of collecting data from parents of non-identified students helps to create a representative sample of the population of U.S. children and adolescents. Collecting data from parents of identified students helps to create populations typically served by groups or organizations that differ substantially from the general sample. Using non-identified and identified students in the study involves comparison to a general representative sample of children and adolescents, which can be important for determining how common or uncommon an obtained score may be and helpful when evaluating how well the BASC-2 BESS Total Score can differentiate between these groups (Kamphaus & Reynolds, 2007). In addition to looking at the correlation between the two instruments, the study also examined if the behavior screener predicted the same group of students that get referred to the office as the behavior rating scale.

Consistency between the two measures is important. The BASC-2 BESS screener needs to assess for the same information in order to be effective in following up with the BASC-2 as a diagnostic instrument. The hypotheses are:

1. The BASC-2 BESS and BASC-2 PRS will be highly correlated.

2. The subscales of similar constructs will be highly correlated on the BASC-2 BESS and BASC-2 PRS.

3. The BASC-2 BESS and BASC-2 PRS will predict referrals to the office.

Chapter 2

Method

Participants

A total of 54 parents of students of an elementary school in a predominantly rural county in southeast Ohio were sent letters with information about the study. A total of 26 parents (48%) sent back permission forms (with one attempt to secure participation and no follow-up), and ultimately, data were collected from 23 parents (42%) who comprise the two groups in this study. Group one consisted of parents of 8 identified students (35%), ages 6-14 years. Students identified were recruited from the school's special education list of students with IEPs. The identified group included 7 males and 1 female. Group two consisted of parents of 15 non-identified students (65%), ages 7-13 years, recruited with encouragement from the school's principal. The non-identified group included 5 males and 10 females.

Procedure

Parents of both groups were contacted by the principal with a letter describing the purpose and procedures of the study, along with a cover letter written by the principal encouraging parent participation. The principal of the elementary school was responsible for obtaining parental consent. Parents who chose to participate signed a permission form to the principal giving permission for their child's scores to be used in the study. Once the parent

permission form was returned, a packet containing the BASC-2 BESS and BASC-2 PRS were sent home to the parents. The parent who has the most frequent contact with his or her child was requested to do the ratings. Parents were instructed to respond anonymously but were asked to indicate whether they are either a mother, father, or guardian. To maintain confidentiality, the principal created a coding system. Each rating scale was letter coded for gender, numbers were used in place of names, two check marks indicated if the student had more than 3 referrals to the office, and one X indicated if the student is receiving services from an IEP or 504 plan. Students without an X on their data are the non-identified group. All rating scales were scored using the computer scoring software sold by the tests' publishers or hand scored. This study did not involve any incentives and participation was voluntary.

Instruments

BASC-2 Behavioral and Emotional Screening System (BASC-2 BESS). The BASC-2 BESS is designed to determine behavioral and emotional strengths and weaknesses in children and adolescents ages 3 to 18 years. It consists of brief screening measures that can be completed by teachers, parents, and students. The BASC-2 BESS offers numerous features that make it efficient and effective for identifying behavioral and emotional problems in children and adolescents. The features include a single Total Score that is reliable and accurate for predicting a broad range of behavioral, emotional, and academic problems; forms that can be completed in 5 minutes or less; assessment of a wide variety of behaviors that represent both problems and strengths, including internalizing problems, externalizing problems, school problems, and adaptive skills; normative samples that are closely matched to the U.S. population; and scoring services that can generate group-level aggregate reports that may be helpful for tracking progress over several years' time. The BASC-2 BESS consists of Teacher and Parent Forms at the

Preschool level (ages 3 through 5), Teacher and Parent Forms at the Child and Adolescent level (grades K through 12) and Student Forms at the Child and Adolescent level (grades 3 through 12). Each form contains between 25 and 30 items and can be completed in 5 minutes or less. Only the Parent Form Level CA (Child/Adolescent) Grades K through 12 was used in the present study. Items for the BASC-2 BESS originated from the pool of items created during the development of the BASC-2 Teacher Rating Scales, Parent Rating Scales, and Self-Report of Personality. Each item pool consists of over 650 items. These items focused on behavioral and emotional strengths and weaknesses and covered a wide range of areas, including externalizing problems (hyperactivity, aggression), internalizing problems (anxiety, depression), adaptive skills (functional communication, social skills), and school problems (attention problems, learning problems). T scores between 61 and 70 are considered elevated and scores of 71 and above are considered extremely elevated. As reported in the BESS manual, the Child/Adolescent Parent Form evidenced a split-half reliability estimate ranging from .94 to .95 and an adjusted test-retest correlation coefficient of .86. Children and adolescents, aged 3 through 18, who were included in the norm samples were from a total of 233 cities in 40 states and are representative of the general U.S. population in terms of sex, socioeconomic status, race/ethnicity, geographic region, and classification in special-education programs. A total of 12, 350 cases were selected. For the combined norms, the number of cases in an age group ranged from 250 to 700 at the Preschool level and from 700 to 1,600 at the Child/Adolescent level (Kamphaus & Reynolds, 2007).

Behavior Assessment System for Children Second Edition (BASC-2). The BASC-2 is a multi-dimensional assessment system that evaluates clinical and adaptive aspects of behavior and personality. Whereas the BASC-2 includes rating scales that can be completed by the child

and/or teacher, only the Parent Rating Scale (PRS) was used in the present study and will be described here. The PRS provides information about students' problem and adaptive behaviors at home and in the community, and is available for three age ranges including preschool (ages 2-5 years), child (6-11 years), and adolescent (12-21 years). Only the Child (PRS-C; 160 items) and adolescent (PRS-A; 150 items) forms will be used in the present study. In completing the PRS, the parent or guardian is asked to read phrases that describe how children may act and then rate their child's behavior based on the last several months relative to the phrase. Each PRS item is rated on a four-point frequency scale (0 = Never, 1 = Sometimes, 2 = Often, and 3 = Almost Always) and item raw scores are summed and converted into standardized T scores (M = 50; SD = 10) for interpretation. For the clinical scales, high scores represent more problematic behaviors. T scores between 60 and 69 are considered at-risk and scores of 70 and above are considered clinically significant. For the adaptive scales, lower scores represent deficits. T scores between 31 and 40 fall in the at-risk range and scores of 30 and below are considered clinically significant. Different sets of related PRS items form nine clinical scales (Aggression, Anxiety, Attention Problems, Atypicality, Conduct Problems, Depression, Hyperactivity, Somatization, and Withdrawal) and five adaptive scales (Activities of Daily Living, Adaptability, Functional Communication, Leadership, and Social Skills). Together, the nine clinical and five adaptive scales are used to generate four PRS composites: Externalizing Problems, Internalizing Problems, Behavioral Symptoms Index, and Adaptive Skills. BASC-2 PRS psychometric properties are reported in the manual and are considered strong. For the BASC-2 PRS-C and PRS-A forms, internal consistency reliability coefficients for the major composite scores ranged from .90 to .95, while alpha coefficients for the individual scales ranged from .72 to .88. Concurrent validity studies with a variety of child behavior rating scales (ASEBA Child

Behavior Checklist) yielded moderate to high correlations between scales measuring similar constructs. The General combined-sex norm sample is representative of over 5,000 children and adolescents up to 18 years old. The demographics of this sample are closely matched to the 2000 U.S. Census population with respect to race/ethnicity, parent education, geographic region, and clinical or special education classification (Reynolds & Kamphaus, 2004; Knoll et al., 2010).

Chapter 3

Results

Strength of Correlations

Table 1 provides the Pearson correlation coefficient that was calculated for the relationship between the T score of the BASC-2 BESS and most scales on the BASC-2 PRS. A strong positive correlation was found between the BASC-2 BESS and the BASC-2 Externalizing Problems, Behavior Symptoms Index, Hyperactivity, and Attention Problems scales indicating a significant linear relationship between the variables (Externalizing Problems r (21) = .782, p < .001; Behavior Symptoms Index r (21) = .863, p < .001; Hyperactivity r (21) = .760, p < .001; Attention Problems r (21) = .865, p < .001). A moderate positive correlation was found between the BASC-2 BESS and the BASC-2 Internalizing Problems, Aggression, and Conduct Problems scales indicating a significant linear relationship between the variables (Internalizing Problems r (21) = .536, p > .001; Aggression r (21) = .630, p = .001; Conduct Problems r (21) = .660, p = .001). A strong negative correlation was found between the BASC-2 BESS and the BASC-2 Adaptive Scale indicating a significant linear relationship between the two variables (Adaptive Scale (r) 21 = -.919, p < .001).

Hits and Misses

A comparison of hits (accurate prediction of referral) and misses (inaccurate prediction of referral) of each measure helped to determine the predictive validity of each instrument. The BASC-2 scales (T score 60 and above) and the BESS T score (61 and above) were used to identify the accurate prediction of referrals. Those same BASC-2 scales (T score 59 and below) and the BESS T score (60 and below) were also used to identify the inaccurate prediction of referrals. Students who were identified by the scales who did not receive office referrals are described as false positives. Table 2 provides the comparison of hits and misses of each instrument. According to the data, the BASC-2 BESS accurately predicted 5 of the 6 referrals to the office while the BASC-2 accurately predicted 4 of the 6 referrals to the office. The BASC- 2 BESS inaccurately predicted 1 referral to the office while the BASC-2 inaccurately predicted 2 referrals to the office. The BASC-2 BESS and BASC-2 each had 1 false positive.

Chapter 4

Discussion

Conclusions

The current study compared the parent versions of the BASC-2 BESS and the BASC-2 to determine the congruent validity between the instruments in a group of identified and non-identified students. Predictive validity was examined by comparing the instruments to office referrals. A strength of this study is that both instruments were completed at the same time by the same raters. Any differences in results cannot be due to different raters, settings, or time. This study provided the needed research information regarding the relationship between the BASC-2 BESS and the BASC-2.

Correlations between the BASC-2 BESS and the subscales of the BASC-2 were statistically significant and generally strong, providing evidence that the BASC-2 BESS Total score measures a variety of behavioral and emotional problems and supports the use of the BASC-2 BESS as a screener. Items for the BASC-2 BESS originated from the pool of items created during the development of the BASC-2 Rating Scales, and these items covered a wide range of areas, including externalizing problems, internalizing problems, adaptive skills, and school problems (Kamphaus & Reynolds, 2007). The current study found a strong correlation between the BASC-2 BESS Total score and the subscales measuring externalizing behavior and a moderate correlation between the BASC-2 BESS Total score and the BASC-2 subscales measuring internalizing problems. This result suggests that the pool of items used in the development of the BASC-2 BESS included more externalizing behavior than internalizing behavior items. These results are not consistent with the BASC-2 BESS manual which found that a strong relationship exists between the BASC-2 BESS Total Score and all the composite scores available on the BASC-2 (Kamphaus & Reynolds, 2007). This difference may be due to the small sample size. Another study with a larger sample size is needed to evaluate this difference. If the BASC-2 BESS is more sensitive to externalizing behavior problems then this may impact how the instrument is used.

The analysis of hits and misses between the two instruments provided evidence for the BASC-2 BESS being accurate for predicting office referrals and usable for screening students for possible behavioral problems. In this study the BASC-2 BESS was better at predicting referrals to the office. This result may be because the pool of items used in its development included more items from the externalizing than internalizing behavioral composites, and those are the behaviors that get students referred to the office. This outcome may be a possible weakness of

the BASC-2 BESS because often the internalizing behavior problems are more difficult to detect. Teachers are better at identifying children who are aggressive and hyperactive yet may need help in determining the ones who are anxious or depressed.

Limitations and Future Research

One limitation of the current study is that the sample was drawn from a predominantly rural county in southeast Ohio, so it is unknown if the same results would be obtained if parents of other ethnic backgrounds were used with these instruments. This limitation makes it more difficult to generalize to the population as a whole. For future research, exploration of group differences across race/ethnicity is needed to provide further information about the generalizability of the findings. While the intent was to compare identified and non-identified students, the sample was too small to enable the researcher to compare these two groups. Additional research is needed to determine whether the findings from this study apply equally to special education as to regular education students. Given that the results of the BASC-2 BESS' predictive validity may be due to its bias in externalized behaviors, a future study may want to compare two behavior screeners in predicting office referrals, the BASC-2 BESS and the Behavior Intervention Monitoring Assessment System (BIMAS) is one possible comparison. This study only included ratings by parents. Although this information is important, teachers are often the ones called upon to rate the behaviors of their students. A follow up study should compare profiles from different sources including teachers.

Implications for Practice

The current study provides important information about the consistency of scores from BASC-2 BESS and the BASC-2. The BASC-2 BESS is accurate for predicting office referrals;

however, it did not correlate strongly with the internalizing composite from the BASC-2. It is hoped that the information from this study will help school-based practitioners with selecting the most appropriate assessment based on primary purpose of the assessment and the adequacy of the psychometric property of the assessment.

Appendix A

Letter from Institutional Research Board

MARSHALL UNIVERSITY®
w w w . m a r s h a l l . e d u

Office of Research Integrity   FWA 00002704
Institutional Review Board
401 11th St., Suite 1300   IRB1 #00002205
Huntington, WV 25701   IRB2 #00003206

March 28, 2013

Sandra Stroebel
Psychology Department

RE: IRBNet ID# 412647-1
At: Marshall University Institutional Review Board #2 (Social/Behavioral)

Dear Dr. Stroebel:

Protocol Title:	[412647-1] Examining the Consistency of Two Behavior Rating Scales with Identified and Non-Identified Elementary Students

Expiration Date:	March 28, 2014	
Site Location:	MUGC	
Submission Type:	New Project	APPROVED
Review Type:	Exempt Review	

In accordance with 45CFR46.101(b)(4), the above study and informed consent were granted Exempted approval today by the Marshall University Institutional Review Board #2 (Social/Behavioral) Designee for the period of 12 months. The approval will expire March 28, 2014. A continuing review request for this study must be submitted no later than 30 days prior to the expiration date.

This study is for student Amber Wallbrown.

If you have any questions, please contact the Marshall University Institutional Review Board #2 (Social/ Behavioral) Coordinator Michelle Woomer, B.A., M.S at (304) 696-4308 or woomer3@marshall.edu. Please include your study title and reference number in all correspondence with this office.

Appendix B

Table 1

Pearson Correlation Coefficient

BASC-2 Scales								
N = 23	1	2	3	4	5	6	7	8
BESS Pearson Correlation	.782**	.536**	.863**	-.919**	.760**	.630**	.660**	.865**
Significance	.000	.008	.000	.000	.000	.001	.001	.000

Notes. 1 = Externalizing Problems; 2 = Internalizing Problems; 3 = Behavioral Symptoms Index; 4 = Adaptive Skills; 5 = Hyperactivity; 6 = Aggression; 7 = Conduct Problems; 8 = Attention Problems
**Correlation is significant at the 0.01 level

Appendix C

Table 2

BASC-2/BESS Hits and Misses

N = 23	Hits	Misses	False Positives
BASC-2 BESS	5	1	1
BASC-2	4	2	1

References

Angello, L.M., DiPerna, J.C., Gureasko-Moore, D.P., Gureasko-Moore, S.P., Nebrig, M.R., Ota,

    K., & Volpe, R.J. (2003). Assessment of attention-deficit/hyperactivity disorder: An

    evaluation of six published rating scales. *School Psychology Review, 32*, 241-262.

Barber, C.R., Richardson, R.C., Schultz, B.L., &Wilcox, D. (2011). A preschool pilot study

    of connecting with others: Lessons for teaching social and emotional competence. *Early*

    Childhood Educ J, 39, 143-148. doi: 10.1007/s10643-011 0450-4

Bergeron, R., Farmer, W.L., Floyd, R.G., & McCormack, A.C. (2008). The generalizability of

    externalizing behavior composites and subscale scores across time, rater, and instrument.

    School Psychology Review, 37, 91-108.

Blackburn, L.B., Sanders, S.J., Kanive, R., & Titus, J.B. (2008). Behavioral profiles of

    children with epilepsy: Parent and teacher reports of emotional, behavioral, and

    educational concerns on the basc-2. *Psychology in the Schools, 45*, 893-904. doi:

    10.1002/pits.20333

Blaha, J., Merydith, S.P., & Prout, H.T. (2003). Social desirability and behavior rating scales:

    An exploratory study with the child behavior checklist/4-18. *Psychology in the Schools,*

    40, 225-235. doi: 10.1002/pits.10077

Boelter, E.W., Caldarella, P., Gentry, A., Merrell, K.W., & Streeter, A.L. (2001). Validity of the

    home and community social behavior scales: Comparison with five behavior rating

    scales. *Psychology in the Schools, 38*, 313-325.

Bour, J.L., Hakman, M., Murphy, S.B., Myers, C.L., & Sidebottom, K.J. (2010). Same constructs, different results: Examining the consistency of two behavior rating scales with referred preschoolers. *Psychology in the Schools, 47*, 205-216. doi: 10.1002/pits.20465

Burlingame, G.M., Eggett, D.L., McClendon, R.J., McClendon, D.T., Warren, J.S., & Green, K.M. (2011). Sensitivity to change of youth treatment outcome measures:A comparison of the cbcl, basc-2, and y-oq. *Journal of Clinical Psychology, 67*, 111-125. doi: 10.1002/jclp.20746

Chafouleas, S.M., Kilgus, S.P., Maggin, D.M., & Sanetti, L.M.H. (2012). Evaluating sensitivity to behavioral change using direct behavior rating single-item scales. *Exceptional Children, 78*, 491-505.

Chafouleas, S.M., Kilgus, S.P., Riley-Tillman, T.C., & Welsh, M.E. (2012). Direct behavior rating scales as screeners: A preliminary investigation of diagnostic accuracy in elementary school. *School Psychology Quarterly, 27*, 41-50. doi: 10.1037/a0027150

Cook, C.R., Elliot, S.N., Gresham, F.M., Kettler, R., & Vance, M.J. (2010). Cross informant agreement for ratings for social skill and problem behavior ratings: An investigation of the social skills improvement system-rating scale. *Psychology Assessment, 22*, 157-166. doi: 10.1037/a0018124

Gadow, K.D. & Volpe, R.J. (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.

Hawkins, J.D., Catalano, R.F., Kosterman, R., Abbott, R., & Hill, K.G. (1999). Preventing

    adolescent health-risk behaviors by strengthening protection during childhood. *Archives*

    of Pediatrics & Adolescent Medicine, 153.

Hosp, J.L., Hosp, M.K., & Howell, K.W. (2003). Characteristics of behavior rating scales.

    Journal of Positive Behavior Interventions, 5, 201-208.

Kamphaus, R.W. & Reynolds, C.R. (2007). Behavior & emotional screening system

    manual. Bloomington, MN: Pearson.

Knoll, V.A., Lopata, C., Rodgers, J.D., Smerbeck, A.M., Thomeer, M.L., Toomey, J.A., &

    Volker, M.A. (2010). Basc-2 prs profiles for students with high-functioning autism

    spectrum disorders. *J Autism Dev Discord, 40*, 188-199. doi: 10.1007/s10803-009-0849

    6

Reynolds, C.R., & Kamphaus, R.W. (2004). Behavior assessment system for children second

    edition. Circle Pines, MN: AGS.