Georgia State University
Digital Archive @ GSU

Counseling and Psychological Services Dissertations

Department of Counseling and Psychological Services

5-5-2012

Universal Screening as the Great Equalizer: Eliminating Disproportionality in Special Education Referrals

Tara C. Raines
Georgia State University, traines@student.gsu.edu

Follow this and additional works at: http://digitalarchive.gsu.edu/cps_diss

Recommended Citation
Raines, Tara C., "Universal Screening as the Great Equalizer: Eliminating Disproportionality in Special Education Referrals" (2012). *Counseling and Psychological Services Dissertations.* Paper 83.

This Dissertation is brought to you for free and open access by the Department of Counseling and Psychological Services at Digital Archive @ GSU. It has been accepted for inclusion in Counseling and Psychological Services Dissertations by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

ACCEPTANCE

This dissertation, UNIVERSAL SCREENING AS THE GREAT EQUALIZER: ELIMINATING DISPROPORTIONALITY IN SPECIAL EDUCATION REFERRALS, by TARA C. RAINES, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree Doctor of Philosophy in the College of Education, Georgia State University. The Dissertation Advisory Committee and the student's Department Chair, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty. The Dean of the College of Education concurs.

_____
R. W. Kamphaus , Ph.D.
Committee Co-Chair

_____
Bridget V. Dever, Ph.D.
Committee Co-Chair

_____
Andrew T. Roach, Ph.D.
Committee Member

_____
Miles A. Irving, Ph.D.
Committee Member

_____
Gwendolyn T. Benson, Ph. D.
Committee Member

_____
Date

_____
Brian Dew, Ph.D.
Chair, Department of Counseling and Psychological Services

_____
R.W. Kamphaus, Ph.D.
Dean and Distinguished Research Professor
College of Education

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education's director of graduate studies and research, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves financial gain will not be allowed without my written permission.

_____

Tara C. Raines

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Tara C. Raines
8731 Hammond Forest Dr.
Jacksonville, Fl 32221

The director of this dissertation is:

R. W. Kamphaus, Ph.D.
Office of the Dean
College of Education
Georgia State University
P. O. Box 3980
Atlanta, GA 30302-3980

Tara C. Raines

TaraCRaines@gmail.com
8731 Hammond Forest Dr.
Jacksonville, Fl 32221
(904) 382-8020

EDUCATION

Ph.D. **Georgia State University –** Atlanta, GA
School Psychology, May 2012

Psy.S. **Nova Southeastern University –** Davie, FL
School Psychology, May 2007

M.S. **Nova Southeastern University –** Davie, FL
School Psychology, May 2006

M.S. **The Florida State University –** Tallahassee, FL
Special Education, December 2002

B.A. **The Florida State University –** Tallahassee, FL
Special Education, December 2002

PROFESSIONAL EXPERIENCE

July 2011-
Present **APA Pre-doctoral Psychology Intern**
Children's Hospital Los Angeles Los Angeles, CA

February 2008-
Present **Ecuador Professional Preparation Program Co-Founder & Executive Director-**
Quito, Ecuador

July 2007-
June 2011 **School Psychologist**
Gwinnett County Public Schools, Suwanee, GA

August 2006-
June 2007 **School Psychology Internship Experience**
Gwinnett County Public Schools, (K-12), Suwanee, GA

January 2003-
May 2006 **Special Education Teacher**
Paxon Middle School, Jacksonville, Florida

SCHOLARLY PRESENTATIONS AND PUBLICATIONS

Dever, B.V., Kamphaus, R.W., & **Raines, T.C**. (February 2012). *The measurement equivalence of a screener for behavioral and emotional risk across language form.* Paper accepted to the Society for Research in Child Development (SRCD) Developmental Methodology themed meeting, Tampa, FL.

Reynolds, C.R., Kamphaus, R.W., & **Raines, T. C.** (2011). Reynolds Intellectual Assessment Scale (RIAS). In D. Flanagan & P. Harrison (3rd Ed) (Eds.). *Contemporary intellectual assessment* (pp. 23-38). New York: The Guilford Press.

Whisenhunt, J., Chang, C., Flowers, L., Brack, G., O'Hara, C., & **Raines, T.C.** (2011). Working with clients who self-injure: A grounded theory approach. *Journal of Counseling and Development.*

Kamphaus, R. W. & **Raines, T.C**., (2011). Comorbidity and Cumulative Effects of Inattention,

Poor Emotional Control, and Language Problems on Academic Achievement in Early Childhood. In N. Mather & L. Jaffe (Eds.), *Comprehensive Evaluations: Case Reports for Psychologists, Diagnosticians, and Special Educators* New Jersey: Wiley & Sons Publishers.

Raines, T.C. & Reynolds, C. R., (2011). Re-establishing Eligibility and Re-evaluation Post Natural Disaster. In N. Mather & L. Jaffe (Eds.), *Comprehensive Evaluations: Case Reports for Psychologists, Diagnosticians, and Special Educators.* New Jersey: Wiley & Sons Publishers.

Berzins, A.R., & **Raines, T.C.,** (2010). The Influence of Cultural Immersion Programs in Creating Culturally Competent Education Professionals. In K. Warnick, P. Warnick, & A. Laffoon (Eds.), *Educational Policy and Practice: The Good, the Bad and the Pseudoscience Volumes I & II.* New York: Nova Science Publishers.

Raines, T.C. (November 2011) *Homework Made Easy...Strategies for Structuring Afterschool Time*. Workshop presented at the Compton Unified School District 26th Annual District Advisory Committee Parent Training Institute, Ontario, CA.

Fernandez, L., **Raines, T. C**., & Berzins, A.R. (August 2011). *Reflections on Models of Consultation Used in a Cultural Emersion Program: Ecuador.* Poster presented at the 119th Annual Conference of the American Psychological Association in Washington, D.C.

Kamphaus, R.W., Dever, B., **Raines, T. C.,** Dowdy, E., & DiStefano, C. (August 2011). *Implementation of an Adolescent Behavioral and Emotional Risk Surveillance Protocol in Schools*. Poster presented at the 119th Annual Conference of the American Psychological Association in Washington, D.C.

Berzins, A., **Raines, T.C**., Oakland, T., & Dejud, C (February 2011). *The Influence of Cultural Immersion Programs on Cultural Competence.* Symposium presented at the Annual Convention of National Association of School Psychologists in San Francisco, CA.

Mays, K., **Raines, T.C**., & Dever, B (February 2011). *Universal Screening for Behavioral and Emotional Risks in Schools.* Mini-Skills presented at the Annual Convention of National Association of School Psychologists in San Francisco, CA.

Raines, T.C., Crumpler, Y.M, Smith, O.W., & Torres-Villa, M.S. (February 2011*). Development of Bilingual Endorsement for School Psychologists in Georgia,* Participant Information Exchange/Poster presentation presented at the Annual Convention of National Association of School Psychologists in San Francisco, CA.

Flowers, L., Whisenhunt, J., Grubbs, N., **Raines, T.C.,** Shelton, L., Lokesmoe, K., Kargar, N., Dispenza, F., & Brack, G. L. (October 2011). *Posttraumatic Growth with Counselors in Training*. Paper presented at Southern Association for Counselor Education and Supervision, Williamsburg, VA.

Raines, T.C., Juechter, J., Mays, K., & Robertson, A. (March 2010). *Screening for Children at Risk for Behavioral and Emotional Problems.* Poster presented at the Annual Convention of National Association of School Psychologists in Chicago, IL.

Berzins, A.R., Romero, P., & **Raines, T.C**. (March 2010). *Reflections of and Professional Experiences in Quito, Ecuador.* Poster presented at the Annual National Convention of NASP in Chicago, IL.

Ortiz, A., Guimares, M., & **Raines, T.C**. (August 2009). *Pre-Kindergarten Screener Emphasizing a Response-to-Intervention Approach in Ecuador*. Paper presented at the Annual National Convention of American Psychological Association in Boston, MA.

Berzins, A.R., **Raines, T.C.,** & Romero, P. (February 2009). *Ecuadorian Professional Preparation Program: Reflection of Experiences.* Poster presented at the Fortieth Annual National Convention of National Association of School Psychologists in Boston, MA.

ABSTRACT
UNIVERSAL SCREENING AS THE GREAT EQUALIZER:
ELIMINATING DISPROPORTIONALITY IN
SPECIAL EDUCATION REFERRALS
by
Tara C. Raines

The overrepresentation of minority students identified for special education services continues to plague schools and serves as a challenge for researchers and practitioners (Ferri&Conner,2005). Teacher nomination, office discipline referrals (ODR), and functional behavior assessments (FBA) continue to guide referral processes (Bradshaw, Mitchell, O'Brennen, & Leaf, 2010; Eklund, et al., 2009; Mustian, 2010). These methods have been found to be riddled with inconsistencies. Practices used to identify students for behavioral and emotional interventions over-identify students from culturally and linguistically diverse backgrounds. The use of a behavioral and emotional screener to make data-based decisions regarding placement and services could provide an objective assessment of student risk.

The first chapter of this dissertation reviews methods used in the identification of students for behavioral and emotional support services. Additionally, the use of universal screening in conjunction with student self-report are proposed as tools for alleviating the overrepresentation of minority students in special education programs for behavioral and emotional disorders.

The second chapter of this dissertation explores the measurement equivalence of Behavior Assessment System for Children, Second Edition (BASC-2) Behavioral and Emotional Screening System Student form (BESS Student) across the Black, Hispanic, and White participants in the norming sample. The BESS Student as a universal screening tool is poised to alleviate the disproportionate number of children of color

identified by schools as having behavior and emotional disorders. This instrument also provides an avenue to identify students with internalizing disorders who are often overlooked in present referral practices (Bradshaw, Buckley, & Ialongo, 2008; Kataoka, Zhang, & Wells, 2002).

The findings of the measurement equivalence study suggests that the BESS Student is, as designed, identifying behavioral and emotional risk across each of the three groups explored. These findings support the use of a universal screening measure as the first step in a multi-step identification and intervention process. Following up with additional assessment to evaluate the specific areas of risk warranting intervention is pivotal to providing appropriate support services and promoting the behavioral and emotional health of students. Implications for research and practice are discussed.

UNIVERSAL SCREENING AS THE GREAT EQUALIZER: ELIMINATING DISPROPORTIONALITY IN SPECIAL EDUCATION REFERRALS

by

Tara C. Raines

A Dissertation
Presented in Partial Fulfillment of Requirements for the
Degree of
Doctor of Philosophy
In
School Psychology
in
the Department of Counseling and Psychological Services
in
the College of Education
Georgia State University

Atlanta, GA
2011

Copyright by
Tara C. Raines
2011

ACKNOWLEDGEMENTS

Many thanks to my committee members, Dr. R.W. Kamphaus, Dr. Bridget Dever, Dr. Andrew Roach, Dr. Miles Irving, and Dr. Gwen Benson for their patience, insights, encouragement, and guidance. I am appreciative of my research team members who provided innumerable ideas and opportunities to sit in the presence of brilliance.

Ten thousand thanks to my parents, Thomas and Mozella, and my sister, Teresa, who rallied together to see me over this hump. Thank you Bridget & Jill for your dedication to Kami and her fare, your love of coupons, your passion for Bon Temps and its residents, and your endurance through my many peaks and valleys. Thank you Eric & Brandy for your commitment to my success and your willingness to barter when I was in need. Thank you Shannon for your endless cheerleading, belief in my abilities, and mystical manifestation powers. You continue to amaze me. Thank you Robin, Tom, Maria, Dale, and all of my GCPS family who supported me as I juggled many responsibilities and believed I had the capacity to do so. My paperwork never would have made it without your diligence and kindness, Ms. Jantosciak. Thank you, Veronica for your mathematical genius, guidance, and services when I was lost. Most importantly, I would be lost without you, Raymond, who truly scooped and carried me through this process steadfastly providing me Maslow's most basic needs. Thank you for stepping up to all of the domestic tasks that I lacked the cognitive resources to complete. Thank you for loving me.

TABLE OF CONTENTS

Page

List of Tables iv
List of Figures v
Abbreviations.....vi

Chapter
1 SCREENING AS THE GREAT EQUALIZER: ELMINATING DISPROPORTIONALITY IN SPECIAL EDUCATION REFERRALS.....1

Introduction.....1

Disproportionality.....2

Current Identification Practices.....4

Universal Screening……..……10

Benefits of Universal Screening for Behavioral and Emotional Risk……..…….14

Choice of Informants……..……...16

Screening Diverse Student Populations……..……..……20

Implications for Practitioners……..……..22

Implications for Researchers ……...……..23

References.....25

2 MEASUREMENT EQUIVALENCE OF BASC-2 BEHAVIORAL AND EMOTIONAL SCREENING SYSTEM BY RACIAL GROUPS 34

Introduction.....34

Methods.....42

Results.....48

Discussion 56

References.....62

LIST OF TABLES

Table Page

1 Demographics from Norm Sample race and gender.....42

2 BESS Student Items and Factor - Dowdy et al (2011) CFA.....46

3 Model Fit Indices for Confirmatory Factor Analysis Models.....55

4 MANOVA Descriptive Statistics.....57

LIST OF FIGURES

Figure Page

1 Path diagram of final Exploratory Factor Analysis Model (Dowdy et al, 2011)...45

2 Confirmatory Factor Analysis Model 1.....49

3 CFA Model 10 for students identified as Hispanic for the national norming sample of the BASC-2 BESS.....54

4 CFA Model 10 for students identified as White for the national norming sample of the BASC-2 BESS.....54

5 CFA Model 10 for students identified as Black for the national norming sample of the BASC-2 BESS.....55

ABBREVATIONS

AERA	American Educational Research Association
ASEBA	Achenbach System of Empirically Based Assessment
APA	American Psychological Association
BASC-2	Behavior Assessment System for Children, Second Edition
BER	Behavioral and Emotional Risk
BESS	Behavioral and Emotional Screening System
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CLD	Culturally and Linguistically Diverse
EBD	Emotional Behavioral Disorder
EFA	Exploratory Factor Analysis
ESI	Emotional Symptoms Index
FBA	Functional Behavior Assessments
IDEA	Individuals with Disabilities Education Act
NCME	National Council on Measurement in Education
NPV	Negative Predictive Value
ODR	Office Discipline Referral
PPV	Positive Predictive Value
RMSEA	Root Mean Square Error of Approximation
RtI	Response to Intervention
SES	Socio-economic Status
SRMR	Standardized Root Mean square Residual
SSBD	Systematic Screening for Behavior Disorders
SWPBS	School-wide Positive Behavior Support

CHAPTER 1

UNIVERSAL SCREENING AS THE GREAT EQUALIZER: ELIMINATING DISPROPORTIONALITY IN SPECIAL EDUCATION REFERRALS

Introduction

The value of early intervention and prevention programs is generally acknowledged (Blair & Diamond, 2008; Greenberg Domitrovich, & Bumbarger, 2000). Early intervention and prevention programs have been linked to positive school outcomes such as high school completion, promoting increased well-being, and enhanced resilience (Blair & Diamond 2008; Greenberg 2000). Blair and Diamond (2008), for example, found that intervening to improve emotional and behavioral regulation in students at risk for school failure positively impacts their likelihood of academic success. On the other hand, current practices for identifying students in need of behavioral and emotional support in schools often fail to identify all students who need support, are implemented after student problems have increased in magnitude, and tend to identify a large number of minority students (Ferri & Conner, 2005; Skiba, Simmons, Ritter, Kohler, Henderson, & Wu, 2006).

Increasing pressures on school districts and state education agencies to address the disproportionate number of minority students in special education programs have had little impact on the practices employed for identifying students for these programs (Artiles & Bal, 2008; Artiles, Bal & King-Thorius, 2010).School districts continue to use teacher lead referral practices that identify an excessive number of minority students

(Ferri & Conner, 2005; Semmel, Gerber & MacMillian, 1994).These practices also overlook the value of individual child data as these data are rarely used in special education decision-making, including the eligibility determination process (Kim & Rowe, 2004; Ferri & Conner, 2005). For these reasons, a change in the methods used to identify students with behavioral and emotional disorders is warranted. This paper reviews research suggesting that the use of student self-report universal screening instruments may diminish the overrepresentation of students of color in special education programs, and guide early intervention for students at risk for behavioral and emotional disorders.

Disproportionality

Researchers have attempted to understand the cause of, and develop remedies for, overrepresentation of culturally and linguistically diverse students in special education programs for more than five decades (Dunn, 1968; Ferri & Connor, 2005; Ferri, Connor & Connor, 2010; Harris et al, 2004). According to the U.S. Department of Education (2006), minority students are identified with a disability and placed in special education programs at a significantly higher rate than their white peers. In addition, Ferri and Conner (2005) found that students of color are more likely to be placed in special education programs that lead to more restrictive school environments. Their findings suggested that 70% of students labeled as emotionally disturbed and 82 % of students labeled mentally retarded spend more than 21% of their time in school outside of the general education classroom. Hosp and Reschly (2003), use meta-analysis, found that African American students were significantly more likely to be both referred for special education services and found eligible than their White and Hispanic peers.

This overrepresentation and restrictive school placements are also troubling given that research has demonstrated that students identified for special education services may suffer social isolation, lower self-esteem, substandard education, and they are twice as likely to drop out of high school (Deninger, 2008; Harris et al , 2004; Terras, Thompson, & Minnis, 2009; Thurlow, Sinclair, & Johnson, 2002; Waitoller, Artiles, and Cheney, 2010). Students in special education are often exposed to a less rigorous curriculum, instructed at a slower pace, and held to lower academic expectations than their peers. For students who are erroneously identified as requiring special education services, these curriculum inadequacies lead inevitably to poorer academic outcomes (Deninger, 2008).

The school incompletion or dropout rates for these misidentified students is particularly troubling. Students who have not completed high school have high rates of unemployment, have lower salaries if employed, are more likely to need public assistance, and to become involved with the criminal justice system (Ferri & Conner, 2005). Other research suggests that only 20% of students in special education with emotional and behavioral disorders pursue any type of post-secondary education (Wagner et al., 2005). It could be concluded that the practice of placing a disproportionate number of minority students in special education classes places them on a trajectory for diminished life opportunities.

In response to persistent findings of disproportionality the U.S. government placed provisions in both the 1997 and 2004 reauthorizations of the Individuals with Disabilities Education Act (IDEA) aimed at reducing the overrepresentation problem (IDEA, 2004). Despite these efforts, there has been no apparent decrease or change in

special education placement practices since their initiation (Ferri & Conner, 2005; Samuels, 2005).

Based on these findings, it could be argued that numerous changes need to be made in school-based special education identification practices, including provision of early intervention or additional support in the general education setting prior to referral for special education services. Such early intervention services are based on the premise that the effects of poor general education instruction can be mitigated but the effects of disabilities based on severe physiological impairments (i.e. neurological processing deficits) cannot. This premise, however, still does not hold. Some research has found that students who receive inadequate instruction, particularly in fundamental areas like reading, are more at-risk for being wrongly identified later as requiring special education services (Harris et al, 2004). Thus, fundamental change is needed in the identification and prevention models currently used by U.S. school districts.

Current Identification Practices

In their classic study of special education decisions, Ysseldyke and Algozzine (1982) found decisions about placement in special education classes are more dependent on social categorizations (gender, SES, race, etc.) than pupil performance data. They also found that special education placement was predicted primarily by whether or not the child has been referred by a teacher for a suspected disability. Inevitably, referral by a teacher led to special education placement (Algozzine & Ysseldyke, 1982). Their research has been subsequently replicated and expanded since their initial findings, yielding the same results (Gartner & Kerzner, Lipsky; 1987; Gottlieb, Alter, Gottlieb, Wishner, and Yoshida, 1990; Gottlieb & Alter 2004; Klinger & Harry 2006). The current

teacher referral process is idiosyncratic and fraught with inaccuracy, specifically for children suspected of having behavioral and emotional disorders. Skiba and colleagues (1993) described teacher referral practices as an "economic process" in which student performance as well as the variety of classroom resources must be considered. Their research found teacher resources, behavior management strategies, and political climate of the building (i.e. administration discouraging special education referrals) all influencing referral practices. They also found that classroom behavior and academic engagement influenced teacher perception of disability and likelihood of special education referrals. These markers of disability are particularly troubling in light of evidence that students from low SES backgrounds and minority group members vary in their presentation of early behavioral and academic skills due to expectations of varying cultural norms (Prince & Lawrence; 1993).

Teacher referral decisions about child behavioral and emotional problems frequently do not agree with referrals using structured and/or standardized ratings of a children's behavioral and emotional adjustment (Eklund, et al., 2009). Waiting for teacher referral often, unfortunately, results in substantial manifestation of behavioral and emotional problems before intervention services are rendered (Eklund et al, 2009; Feil & Walker, 1995). Many teachers lack specific training in how to identify students with emotional and behavioral problems. Teachers also exhibit problems in the selection and use of evidence-based intervention practices for children in general, and in particular, for children who manifest behavioral and emotional difficulties (Lewis, Hudson, Richter, & Johnson, 2004; Maccini & Gagnon, 2006; Mooney et al., 2004).

Numerous studies have shown that teacher referral for special education is related to the level of student disruption in the classroom. Students with more disruptive behaviors, despite the cause or nature of the disruption, are more likely to be referred for special education (Bradshaw, Buckley, & Ialongo, 2008; Walker, Cheney, Stage, & Blum 2005). Boys are more likely to vary from the "norm" of good student behavior and, therefore, be referred for special education. Mirkin (1982) found that even when academic weaknesses are evident, girls were less likely to receive referral for services because of generally compliant behavior (Anderson, 1997; Mirkin, 1982; Serbin Marchessault, McAffer, Peters, & Schwartzman, 1993; Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). Pas, Bradshaw, and Hershfeldt (2010) found that regardless of individual student risk, students placed in classrooms with teachers who had overall high rates of special education referrals were much more likely to be referred to special education. These findings further substantiate the notion that teacher referral alone may not be the most effective method for selecting students with behavioral and emotional problems.

Teacher referrals are often supplemented with anecdotal and archival data sources such as office discipline referrals (ODRs) and functional behavior assessments (FBA) for the purposes of identifying students in need of behavioral and emotional interventions and/or special education services. These practices are equally problematic. Bradshaw, Mitchell, O'Brennen, and Leaf (2010) determined that ODRs were influenced by school and teacher expectations, student behaviors, and teacher efficacy. Their findings suggest that ODRs are merely a reflection of the teacher's use of this method as a disciplinary strategy and they may not be predictive of student behavioral need. Furthermore, in this particular study, African American students were more likely to receive ODRs than their

peers. These findings were similar to those found by other researchers (Gregory & Mosely, 2004; Cartledge & Dukes, 2008; Shaw & Braden, 1990). FBAs, in contrast, were found to help mediate and diminish externalizing behaviors in students and altering teacher perception of student need for special education services. When students presented less externalizing behaviors, teachers were less inclined to believe they were disabled. Additionally, the use of interventions developed based on the "function" of the behavior as assessed by FBAs have been found to have more desirable outcomes. In a study of the use of function-based interventions with African American students, Mustian (2010) found that the use of such interventions could provide valuable information to pre-referral intervention teams and even prevent referral for special education. However, Mustian acknowledges that knowledge of how to properly and effectively utilize FBAs to develop interventions is generally limited to special educators.

New federal mandates that require documentation of student "Response to Intervention" (RtI) as a key element in the identification of students with disabilities (IDEA, 2004) have not been shown to alleviate disproportionality. RtI provides schools with a framework for organizing instruction for struggling students using research-validated procedures and decision-making structures. RtI includes the use of frequent assessments to identify students requiring additional assistance and determining if the current intervention, as implemented, is effective. RtI promotes both differentiated instruction and an ongoing data-based decision-making process (Fuchs & Fuchs, 2006; Tilly, Harken et al., 2008;). There continues to be a lack of consensus on how RtI models are specified and implemented (Fuchs & Fuchs, 2009). Most states use a multi-tiered model that includes a combination of academic screening, academic intervention using

Curriculum-Based Measures (CBM), and standardized assessment to establish educational placement decisions (Fuchs & Deshler, 2007). However, some states have elected to implement a 4tiered model and others 2or 3tiered models. Variance both in the number of tiers as well as what actions are performed at different tiers leads to markedly different referral procedures not only across states but also often within school districts (Berkeley, Bender, Peaster, & Saunders, 2009)

RtI is commonly depicted as a pyramid, which represents the progression of intensity of interventions that at-risk students may receive. In the three-tiered RtI model, which is most commonly adopted, "Tier" provides increased academic or behavioral intervention. Tier 1 generally includes screening and general education "best practices." Tier 2 provides empirically validated interventions for students struggling to make adequate progress with general instruction, and for those identified as having risk for academic or behavioral difficulties as indicated by Tier 1 screening or identification practices. Tier 3 is the most intense level of intervention and often results in placement in special education programs (Fuchs & Fuchs, 2009). Using this method, RtI should theoretically identify 15%to 20% of children at Tier 1to receive a more intense intervention at Tier 2. Roughly, 5% of children in Tier 2 would receive special education services at Tier 3 (Reschley &Ysseldyke, 2002; Tilly, Harken et al., 2008).

Using the RtI model, schools are now charged with providing data to monitor progress on both academic and behavioral interventions. With RtI, behavioral interventions are employed and should be modified periodically depending on student response, similar to a physician modifying a medication dosage. School-based teams are charged with determining whether or not the student is responding adequately to the

intervention. One benefit of this model of identification is that it moves school teams from the common practice of "problem admiration" to attempting some type of intervention (Henderson, 2009). However, the determination of "adequate" progress remains inconsistent, varying often times by teacher, school team, district, and state. Unfortunately, due to the varying interpretations of the definition of "adequate progress" and the nebulous federal definition of Emotional Disturbance, RtI differs minimally from previous practices and is only marginally better than teacher referral alone (Gresham, 2005).

Optimal RtI programs assess the performance of students individually to determine their needs (Harris-Murri, King, & Rostenburg, 2006) and they consider the importance of culturally-sensitive and appropriate interventions through the RtI process when serving diverse student populations. They posit that an ecological approach considering the influence of student culture on their learning and behaviors is vital in the development of optimal interventions. Theoretically, they also hypothesize that the use of culturally-appropriate interventions combined with more consistent RtI practices will ultimately decrease the overrepresentation of minorities in special education programs.

Universal Screening

Screening is defined by Gridley, Mucha, and Hartfield (1995) as "a brief, global, relatively low-cost procedure used to obtain preliminary information about a wide range of behavior for large groups of children" (p. 213). The goal of the administration of a screener is to produce a quick, inexpensive, initial investigation of an issue (Pangano, Cassidy, Little, Murphy, & Jellinek, 2000). For the purposes of this work, universal screening refers to a systematic approach to identifying students who are demonstrating

behavioral and emotional difficulties or the “risk factors” for the development of such problems. The term “risk factors” refers to environmental experiences or influences that occur before the onset of emotional and behavioral disorders that increase the likelihood of the development of these disorders (Loeber, 1990). Screening informs stakeholders of the “statistical chances” of development of a specific pathology, in this case, behavioral and emotional disorders (Skotko, 2011). A universal screening approach would provide information on the level of risk that could be used in the development of educational intervention plans (Dowdy, Mays, Kamphaus, & Reynolds, 2008; Kamphaus, Rowe, Dowdy, & Hendry, 2006).

Preventive interventions such as universal screening are proactive and provided to all students. For this reason, their potential to stigmatize students is limited (Greenberg, 2000). Additionally, universal screening fosters the implementation of interventions and attempts to eliminate prolonged exposure to risk. Based on the assumption that long-term behavioral and academic problems could be averted through early detection and intervention, early identification of students at risk for behavioral and emotional disorders should be high priority for both educators and researchers (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004; Feil & Walker, 1995; Patterson, Reid, & Dishion, 1992).

Implemented as an “intervention,” universal screening increases the potential for students with risk factors to access appropriate mental health services both through school and the community (Husky et al, 2011). In their study of African American students universally screened for suicidal ideation, Brown and Goldstein-Grumet (2009) found African American students in their sample emoted distress associated with suicidal behaviors that may not have been reported though other avenues. With this information

researchers were able to link 62% of these youth to mental health resources in an effort to promote positive coping.

The aforementioned weaknesses in practices used to identify and refer students for behavioral and emotional interventions might be addressed through the use of an instrument designed to screen for behavioral and emotional risk using student self-report. The information gathered from such screening could be used to design interventions and link students to appropriate mental health services. Furthermore, screening for behavioral and emotional risk reduces the potential for the harmful effects of erroneously labeling students as disabled (Greenberg, 2000).Another advantage is that the use of universal screening measures lead to data-based decisions regarding placement and provision of services based a norm-referenced assessment of student risk. The use of such a screener is currently supported by various professional bodies (AERA, APA, & NCME, 1999).

Despite such support, currently only 2% of schools screen all children for emotional and behavioral problems (Romer, & McIntosh, 2005). Barriers to universal screening include the notion that in doing so, the mental health system may be burdened by the overwhelming number of referrals and over identification of students as requiring mental health support (Husky et al, 2009). While a larger number of students may be initially identified as requiring support at the onset of universal screening for behavioral and emotional risk (i.e. BER), these students will ultimately require fewer resources and have a better overall trajectory as a result of early intervention (Jones et al, 2002). However, this approach to intervention may require the reallocation of school district resources and personnel for optimal implementation of an early intervention driven model. There are, however, few uniform protocols and instruments for collecting

universal screening data. Presently, the schools and districts choosing to participate in such practices may collect different information, using various instruments, making it difficult to utilize the data collected systematically and meaningfully (Dowdy, Ritchey, & Kamphaus, 2010). Barriers to universal screening also include issues regarding the use of active vs. passive consent for collecting BER information from students (Gardner, 2011). Screening practices of known reliability and validity are crucial for promoting classroom practices and school services necessary to ensure that all children succeed in school. Schools commonly screen children for other types of disabilities at preschool and older age levels including problems with vision, hearing, speech, cognitive delays, and academic problems associated with specific learning disabilities, but rarely for emotional and behavior problems (Feil & Walker, 1995). However, early identification and intervention for youth with emotional and behavioral problems can help to minimize the long-term detriment of mental disorders (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004; Campaign for Mental Health Reform, 2005)

The use of a universal screening tool is also in alignment with the Response to Intervention model. As previously outlined, this model recommends high-quality instructional practices at tier one and the use of universal screening tools to identify students at risk for disabilities (Fuchs & Fuchs, 2007). The use of a universal screening tool also fits well into the multi-gate/tier approach. With the results of a screening instrument, students can be triaged and receive appropriate interventions based on their level of risk in a timely fashion.

Traditional school practices may delay identification of students who need support until full manifestation of behavioral and emotional problems. Furthermore,

students who demonstrate significant behavioral and emotional problems in childhood are more likely to abuse drugs, drop out of high schools, and develop long-term disabling conditions in adulthood (Conroy & Brown, 2004). With the most commonly utilized identification practices, there is generally a substantial period of time between when the student begins exhibiting symptoms of behavioral and emotional difficulties and when they begin to receive treatment (Duncan, Forness, & Hartsough, 1995). Walker and colleagues (2000) found that referrals for students with academic difficulties were most common in 2nd or 3rd grades. In contrast, referrals for students with behavioral and emotional problems were higher in 9th grade. These findings are disheartening as symptoms of emotional or behavioral problems are often present as early as age 3 (Kazdin, 1987). In part, the problem is due to the current "wait to fail" service delivery mechanisms in which services are not initiated until significant emotional and behavioral concerns are present. This occurs in spite of evidence suggesting that the longer children's behavioral and emotional symptoms go unidentified and untreated, the more stable these symptoms are, making intervention efforts more difficult (Gottlieb, 1991).

Benefits of Universal Screening for Behavioral and Emotional Risk

Through screening, both prevention and intervention work are able to begin simultaneously (Barnett et. al, 2006; Fuchs & Fuchs, 2007). Additionally, there are various options for the implementation of a universal screening program. School-Wide Positive Behavioral Support (SWPBS) programs have been used as universal screening tools. SWPBS programs are prevention-oriented models designed to teach, monitor, and encourage positive school behaviors. This approach is generally multi-tiered and implemented school-wide. Some SWPBS programs have components that allow schools

to measure student risk based on the individual school expectation for behavior. By using the standards set by the SWPBS programs, schools are able to assess which students are demonstrating the behaviors that place them at greatest risk for not being successful in program participation. Simultaneously, the SWPBS program provides both prevention and intervention support for all students (Burke et. al, 2010; Glover & Albers, 2007).

Another method for universal screening for behavioral and emotional risk is the Systematic Screening for Behavior Disorders (SSBD) system (Glover & Albers, 2007) This system, similar to the SWPBS method of screening, is multi-tiered and aims to intervene with students who are not meeting the school expectation for pro-social behavior. The SSBD utilizes teacher nomination of at-risk students (Tier 1) and ratings of nominated students' adaptive and maladaptive behaviors (Tier 2). A structured observation instrument is then applied to students who generate rating scale scores in Tier 2 that suggest substantial risk (Tier 3). In their research of the implementation of the SSBD program, Glover and Albers found schools were able to not only expand the quantity and types of students identified as at risk but they were also able to proactively intervene with these students, providing a wide range of interventions (Glover & Albers, 2007; Walker & Severson, 1990).

Universal screening can also be implemented using a more standardized normative approach. Instruments such as the BASC-2 Behavioral and Emotional Screening System (BASC-2 BESS; Kamphaus & Reynolds, 2008) provide assessment a wide range of behavioral and emotional risk factors compared to a norming sample. In contrast to SWPBS and SSBD approaches, using a standardized measure to assess risk may remove the inconsistencies of teacher nomination from the screening process. In

addition, using rating scales allows for the identification of students with risk for internalizing disorders that may be missed by teacher nomination (Walker, 2005). This would, hopefully, lead to the identification of students with the greatest risk. By accurately identifying students with the greatest risk and providing increasing levels of interventions, the number of students referred for special education may be reduced.

Choice of Informant

Parent and Teacher interviews and surveys, self-report interviews, and self-report surveys are frequently utilized to gather information about behavioral and emotional risk and problems (Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Gross, Fogg, Young, Ridge, Cowell, & Richardson et al, 2006; Jaccard, 1998; Sweeting & West, 1998). Parents are often used as informants to provide information about behavioral functioning (Achenbach, McConaughy, & Howell, 1987; Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Gross, Fogg, Young, Ridge, Cowell, & Richardson et al, 2006; Jaccard, 1998),yet research is conflicting on the role of parents as informants about child behavior. Both the utility and validity of parents as informants has been repeatedly studied empirically (Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Sweeting & West, 1998; Gross, Fogg, Young, Ridge, Cowell, & Richardson et al, 2006.)

Parental reporting of emotional functioning and negative behavior of their children has been found to have the greatest difference from child self-report, suggesting parents may not be most attuned with the social and emotional functioning of their children (Waters, Stewart-Brown, & Fitzpatrick, 2003). In their study of parent report of adolescent depressive symptoms, Moretti (1985) found that the depressive symptoms of the parents impacted their perception of symptoms in their children. Parents who were

experiencing symptoms often saw the same in their children in contrast to the adolescent self-report. When compared to other informants such as teachers and mental health workers, parent report was also found to yield different findings (Moretti, 1985). In contrast, on their study of Dutch students using the Strengths and Difficulties Questionnaire, van Widenfelt, Goedhart, Treffers, and Goodman (2003) found parent and self-report of behavioral and emotional risk to be comparable. Additionally, the findings of Halvorsen, Andersen, and Heyerdahl (2005) in their comparison of parent and self-report of emotional and behavioral functioning in patients with anorexia nervosa also were similar. As a whole, this conflicting research would suggest, as informants, parents may not be the most reliable choice for gathering information about student behavioral functioning.

Because students spend a substantial portion of their day in the school setting, teachers are also frequent reporters for studies attempting to gather information about student behavior (Serbin, Marchessault, McAffer, Peters, & Schwartzman, 1993; Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Lee, Elliot, & Barber, 1994). Similar to parents, research regarding teachers as informants is inconclusive. Researchers have demonstrated that teachers are inconsistent as informants (Epkins & Meyers 1994; Eklund, et al., 2009). However, teacher report of behavior has been found to be more predictive of behavioral and emotional outcomes than parent reports in previous research (Sharp, Croudace, Goodyer, &Amtmann, 2005). Ollendick, Oswald, and Francis (1989) found in their assessment of risk for behavioral and emotional problems, teacher-report yielded similar findings to peer- and self-report of behavior and risk. However, when compared to self-report in Epkins and Meyers (1994), teacher-report of depression in

girls was significantly different. Their study revealed that teachers often overlooked symptoms of this internalizing construct. In all, the findings of these studies suggest that, while information gathered from teachers may not be aligned with self-report, it is valuable information (Sharp, Croudace, Goodyer, & Amtmann, 2005).

As a tool for gathering information about behavioral functioning, self-report is used for innumerable constructs such as substance abuse and emotional functioning (Adams, Kelley, & McCarthy. 1997; Knight, Little, Losoya, & Mulvey, 2004; Lau, McCabe, Yeh, Garland, Hough, & Landsverk, 2003). In all populations, self-reports can be considered a very desirable form of data collection. Self-report presumably reduces bias that may be found in interviewing (e.g. social desirability bias and interviewer bias). Differences between the use of self-report surveys or questionnaires and interviews have been researched (Achenbach, 2006; Blount, Evans, Birch, Warren, & Norton, 2002). Self-report surveys or questionnaires remove interviewer bias and in most cases limit the influence of social desirability in participant responses that is common when information is gathered through interviewer. Additionally, the use of an interview may be time-consuming, as extensive training of interviewers as well as interrater reliability between interviewers must be established. In contrast, self-report surveys or questionnaires are generally low cost and easily distributed. Self-reports are also a valuable and preferred method of gathering information for personality and behavioral data (Blount, Evans, Birch, Warren, & Norton, 2002).

Student self-report of more covert constructs such as behavior and emotional functioning has been found to differ from information obtained from other informants such as parents and teachers (Achenbach, McConaughy, & Howell, 1987, Achenbach,

Krukowski, Dumenci, & Ivanova, 2005; Lam & Cheng, 2002; Sweeting & West, 1998; Waters, 2003). In a meta-analytic review of cross informant correlations for the ASEBA forms, Achenbach and colleagues (1987) found little correlation between child self-report of behavior and adult informant reports when assessing for behavioral and emotional problems. Their findings suggest that situational specificity has great impact on the ratings of children's behavioral and emotional functioning. Environmental structures, demands, and expectations may result in variance in reporting of behavior based on informant. In a later meta-analytic review of literature, Achenbach, Krukowski, Dumenci, and Ivanova (2005) compared the predictive validity of self-report with the predictive validity of reports by parents, caregivers, and teachers and found little correlation between the self-report and informant report. Additionally, the German study using the Screen for Child Anxiety Related Emotional Disorders (SCARED),Weitkamp, Romer, Rosenthal, Wiegand-Grefe, and Daniels (2010) explored self-report of anxiety related disorders in children and adolescents. They discovered moderate agreement between informants and self-report may yielded more symptoms for anxiety disorders than parent report.

Overall, it would appear that information gathered from self-report across ages may differ from information yielded by other informants. However, both the self-report and reports gathered from informants are believed to provide valuable information about behavioral functioning that should be integrated for the purposes of diagnosis, treatment, and program planning. The findings of the studies reviewed regarding informant-report and self-report have repeatedly found differences between self-report and other informants. However, researchers also attest that different informants contribute differing

but valid information, and the outcomes of the self-report and the informant report should be integrated for global outcomes (Achenbach, McConaughy, & Howell, 1987; Weitkamp, Romer, Rosenthal, Wiegand-Grefe, & Daniels, 2010). While screening using multiple informants is may provide a broad range of information regarding student functioning, it is not practical. This approach is likely to be both costly and time consuming. Additionally, while some researchers suggest each informant provides a different perspective, studies have also found that information collected from an additional informant provided little variance in information above and beyond what was provided by the initial informant (Biederman, Keenan, & Faraone, 1990) Pragmatically, in a school context, asking teachers to complete screening questionnaires on each of their students may not be feasible. This approach may be particularly difficult in middle and upper grades where teachers see hundreds of students. For this reason, in addition to the research indicating that youth are less likely to disclose covert mental health constructs to the adults, self-report appears to be the most practical method for collecting universal screening of child behavioral and emotional problems.

Screening diverse student populations

There is support for the use of screening as an approach to early intervention (Jones, et al., 2002; Kamphaus et al., 2007). However, the body of research supporting screening with culturally and linguistically diverse (CLD) populations is sparse. Measurement equivalence of reporting of behavioral and emotional functioning have found that students from CLD backgrounds may indeed yield different ratings from their peers when self-reporting (e.g. Adams, Kelley, & McCarthy. 1997; Knight, Little, Losoya, &Mulvey, 2004;Lau, McCabe, Yeh, Garland, Hough, & Landsverk, 2003).

Adams and colleagues (1997) found in their study of self-reporting of behavioral and emotional difficulties that African American females reported more difficulties in the areas of poor work habits and social problems than their White peers on the Adolescent Behavior Checklist. In an investigation of the measurement equivalence of an instrument to measure self-reporting of juvenile offending, Knight and colleagues (2004) determined that while their instrument generally measured equivalent constructs, there were differences in the reporting patterns between African American and white adolescents, as well as between Hispanic and white adolescents. Measurement equivalence studies using parents and teachers as informants have also yielded results that suggest differences across racial and ethnic groups. (Achenbach, McConaughy, & Howell, 1987; Achenbach & Rescorla, 2000;Gross, Fogg, Young, Ridge, Cowell, & Richardson et al, 2006). These studies, among others support the need for investigation of the measurement equivalence of screening measures to determine their comparability of measurement across CLD groups.

Research using measurement equivalence testing to explore the validity of screening instruments across CLD populations is limited. In their recent publication, Dowdy, Dever, DiStefano, and Chin (2011) explored teacher reporting of BER using the Behavioral Emotional Screening System (BESS) in students with limited English proficiency. Their findings indicated teachers reported students with limited English proficiency to have more learning problems and fewer adaptive skills than their English proficient peers. This research shines light on the need additional research into the measurement equivalence of screening measures across CLD groups. It is imperative

that we thoroughly examine the validity and function of instruments used for screening to promote screening practices that are equitable across groups.

Implications for practitioners

The stakes are high for children classified by school personnel as having emotional or behavioral disorders (EBD). Children with this special education classification are known to have poorer academic achievement and social outcomes, and are twice as likely to drop out of high school (Terras, Thompson, &Minnis, 2009; Thurlow, Sinclair, & Johnson, 2002; Waitoller, Artiles, and Cheney, 2010). Research has also shown that the stakes are higher for children of color, especially boys, because these children are classified as EBD at a rate that is far higher than would be predicted by population proportions (Bradshaw, Mitchell, O'Brennan, & Leaf,2010). In other words, disproportional classification rates for EBD have an even more harmful impact on children of color, their families, and communities.

The use of universal screening tools encourages taking steps to assess risk for disabilities in order to develop interventions and if needed refer for special education services (Brown & Barlow, 2005). Screeners provide an overview of different levels of functioning in various domains and relevant individual student data. In particular, screeners “tap skills believed to be related to school learning tasks that are predictive of school success,” (Gredler, 1997, p. 99). Yet, the use of universal screening tools rarely carries over into elementary, middle, and high schools. Furthermore, Blair and Diamond (2008) found that intervening to improve children’s emotional, attention, and behavioral regulation in students at risk for school failure could improve their likelihood of

academic success. Universal screening may serve as a method for ensuring that all children have equal opportunity to have their academic, social, and behavioral needs met without reliance on the varying judgment of teachers alone.

Current flaws in the special education referral system might be addressed through the use of universal screening to make data-based decisions regarding placement and provision of services. Failure to provide early intervention services results in dire outcomes. Barry (2008) determined that attention problems and delayed behavioral skills, as identified by the BASC-2 Teacher Rating Scale in third grade students were predicative of high school dropouts with 80% accuracy. Blair and Diamond (2008) found that intervening to improve children's emotional, attention, and behavioral regulation in students at risk for school failure improves their likelihood of academic success. Screening may serve as a method for ensuring that all children have equal opportunity to have their behavioral and emotional needs met, while also potentially addressing the problem of overrepresentation. The use of such a screener is currently supported by various professional bodies (AERA, APA, &NCME, 1999).

Implications for research

This proposition raises several subsequent questions to be answered by future research. Longitudinal studies of students identified through universal screening may provide relevant information to guide screening practices. Information regarding the outcomes of universal screening ratings across raters and student race may be beneficial to research on disproportionality. Research in this area would not only support research in disproportionality but also early intervention and screening research.

Research seeking to determine if the nature of teacher perceptions of elevated risk and behavioral and emotional problems are a result of cultural differences between the teachers and the students may also be beneficial. As was previously asserted, a large factor in the special education referral process is dependent on teacher nomination and recommendation. An examination of teacher behavioral and social expectations may assist in the exploration of the impetus behind the elevated ratings and frequent referrals for special education services.

Finally, an evaluation of the measurement equivalence and other psychometric properties of universal screening instruments by student race may provide valuable information about similarities and differences in the functioning of screening tools by student group. It has been found that 20% of the school age students require treatment for some emotional and behavioral difficulty, however, just under 1% of the school population is eligible for special education support for emotional and behavior disorders (Burns & Hoagwood, 2002). The 1% identified is overwhelmingly African American. For these reasons among others illustrated in this work, research suggesting measurement equivalence across diverse groups would suggest that it is imperative that researchers and practitioners embrace the tools of universal screening and self-report.

References

Abebe, S., &Hailemariam, A. (2008). Factors Influencing Teachers' Decisions to Refer Students to Special Education Evaluation.*Online Submission*.

Achenbach, T. (2006). As Others See Us: Clinical and Research Implications of Cross-Informant Correlations for Psychopathology.*Current Directions in Psychological Science*, *15*(2), 94-98. doi:10.1111/j.0963-7214.2006.00414.x.

Achenbach, T., Krukowski, R., Dumenci, L., &Ivanova, M. (2005). Assessment of Adult Psychopathology: Meta-Analyses and Implications of Cross-Informant Correlations. *Psychological Bulletin*, *131*(3), 361-382. doi:10.1037/0033-2909.131.3.361.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent Behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232. doi:10.1037/0033-2909.101.2.213

Achenbach, T., &Rescorla, L. (2000).*Manual for the ASEBA Preschool Forms and Profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Adams, C., Kelley, M., & McCarthy, M. (1997). The Adolescent Behavior Checklist: Development and initial psychometric properties of a self-report measure for adolescents with ADHD. *Journal of Clinical Child Psychology*, *26*(1), 77-86. doi:10.1207/s15374424jccp2601_8.

American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Aos S, Lieb R, Mayfield J, Miller M, Pennucci A: *Benefits and Costs of Prevention and Early Intervention Programs for Youth.* Washington , Washington State Institute for Public Policy; 2004.

Artiles, A., &Bal, A. (2008). The next generation of disproportionality research: Toward a comparative model in the study of equity in ability difference. *The Journal of Special Education*. 42(1), 4-14. Doi:10.1177/0022466907313603

Artiles, A., Bal, A., King-Thorius, K. (2010). Back to the Future: A Critique of Response to Intervention's Social Justice Views. *Theory Into Practice*. 49(4), 250-257.

Barnett, D. W., Elliott, N., Wolsing, L., Bunger, C., Haski, H., McKissick, C., & Vander Meer, C. D. (2006). Response to Intervention for Young Children With Extremely Challenging Behaviors: What It Might Look Like. *School Psychology Review*, 35(4), 568-582. Retrieved from EBSCO*host*.

Biederman, J., Keenan, K., & Faraone, S. (1990). Parent-based diagnosis of attention deficit disorder predicts a diagnosis based on teacher report. Journal of the American Academy of Child & Adolescent Psychiatry. 29(5), 698-701. doi:10.1097/00004583-199009000-00004

Berkeley, S., Bender, W., Peaster, L., & Saunders (2009). Implementation of Response to Intervention: A Snapshot of Progress. *Journal of Learning Disabilities*.42(1), 85-95.

Blair, C.& Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. Development and Psychopathology, 20 (3) 899-911.

Blount, C., Evans, C., Birch, S., Warren, F., & Norton, K. (2002). The properties of self-report research measures: Beyond psychometrics. *Psychology and Psychotherapy: Theory, Research and Practice*, *75*(2), 151-164. doi:10.1348/147608302169616

Bradshaw, C. P., Buckley, J. A., & Ialongo, N. S. (2008). School-based service utilization among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly*, 23(2), 169-186. doi:10.1037/1045-3830.23.2.169.

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *Journal of Educational Psychology*, 102(2), 508-520. doi:10.1037/a0018450.

Brown, M., Grumet-Goldstein, J. (2009). School-based suicide prevention with African American youth in an urban setting. *Professional Psychology: Research And Practice*. 40 (2),111-117

Burns, B. J. & Hoagwood, K. (2002) Community treatment for youth: Evidence-based interventions for severe emotional and behavioral disorders, Oxford University Press, New York.

Campaign for Mental Health Reform. (2005). A public health crisis: Children and adolescents with mental disorders Congressional briefing. Retrieved September 1, 2005, from www.mhneform.org/kids

Conroy, M. A., & Brown, W. H. (2004). Early Identification, Prevention, and Early Intervention with Young Children At Risk for Emotional or Behavioral Disorders: Issues, Trends, and a Call for Action. *Behavioral Disorders*, 29(3), 224-236. Retrieved from EBSCO*host*.

DeMarco, A., Deretich, S. (2006). Contributing Factors in Special Education Referrals for Emotional or Behavioral Problems. *Online Submission*.

Deninger, M., & Massachusetts Department of Education, O.(2008). Disproportionality: A Look at Special Education and Race in the Commonwealth. Education Research Brief. *Massachusetts Department of Elementary and Secondary Education.*

DiStefano, C., & Morgan, G. B. (2010). Evaluation of the BESS TRS-CA using the Rasch Rating Scale model. *School Psychology Quarterly*, 25(4), 202-212. doi:10.1037/a0021509.

Dowdy, E., Dever, B., DiStefano, C., & Chin, J. (2011). Screening for emotional and behavioral risk among students with Limited English Proficiency. School Psychology Quarterly, 26(1), 14-26. doi:10.01097/a0022072

Dowdy, E., Mays, K. L., Kamphaus, R. W., & Reynolds, C. R. (2009).Roles of diagnosis and classification in school psychology. In T. B. Gutkin & C. R. Reynolds (Eds.). The Handbook of School Psychology (4th ed.). New York: John Wiley & Sons.

Dowdy, E., Ritchey, K., & Kamphaus, R.W. (2010) School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health, 2* (4) 166-176. doi:10.1007/s12310-0109036

Dunn, L. M. (1968). Special education for the mildly retarded: Is much of it justifiable? *Exceptional Children*, *35*(1), 5–21.

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

Epkins, C., & Meyers, A. (1994). Assessment of childhood depression, anxiety, and aggression: Convergent and discriminant validity of self-, parent-, teacher-, and peer-report measures. *Journal of Personality Assessment*, *62*(2), 364-381. doi:10.1207/s15327752jpa6202_16.

Feil, E. G., & Walker, H. M. (1995). The Early Screening Project for young children with behavior problems. *Journal of Emotional & Behavioral Disorders*, 3(4), 194. Retrieved from EBSCO*host*. doi:10.1177/106342669500300401

Ferri, B.A., & Connor, D. J. (2005). Tools of Exclusion :Race, disability, and

(re)segregated education. Teachers College Record *107*(3), 453.doi:10.1111/j.1467-9620.2005.00483.x

Ferri, B., Connor, D., & Connor, D. (2010). 'I Was the Special Ed. Girl': Urban Working-classYoung Women of Colour. *Gender & Education*, *22*(1), 105-121. doi:10.1080/09540250802612688

Fuchs, D. and D. D. Deshler (2007). What we need to know about Responsiveness to Intervention (and Shouldn't be afraid to ask).*Learning Disabilities Research and Practice,* 22(2), 129-136.

Fuchs, L. S. and D. Fuchs (2009). On the importance of a unified model of Responsiveness to Intervention. *Child Development Perspectives,* 3(1), 41-43.

Fuchs, L. S., & Fuchs, D. (2006). A model for implementing responsiveness to intervention. *Teaching Exceptional Children*, 39(5), 14-20.

Gardner, N. (2011) *Potentials and Impediments to Universal, School-based Screening for Behavioral and Emotional Risk: A Critical Discourse Analysis of Current Case Law*. Unpublished doctoral dissertation, Georgia State University, Atlanta, GA.

Gartner, A., &Kerzner Lipsky, D., (1987).Beyond Special Education: Toward a Quality System for All Students, *Harvard Educational Review,367* (57), 374-76.

Glover, T., & Albers, C., (2007). Universal Screening for Enhanced Educational and Mental Health Outcomes. *Journal of School Psychology*, *45*(2), 117-135. doi:10.1016/j.jsp.2006.05.005.

Gottlieb, G (1991). Experiential canalization of behavioral development: Theory. Developmental Psychology 27:4–13doi:10.1037//0012-1649.27.1.4

Gottlieb, J., Alter, M., Gottlieb, B., Wishner, J., & Yoshida, R. K. (1990). Final report for year III of the consulting teacher program. Report submitted to New York State Education Department.

Gottlieb, J., & Alter, M. (1994). Special education in urban America: It's not justifiable for many. *Journal of Special Education*, *27*(4), 453. Retrieved from EBSCO*host*. doi:10.1177/002246699402700406.

Greenberg, M.T., Domitrovich, C., &Bumbarger, B. (2000). Preventing mental disorder ins chool-aged children: A review of the effectiveness of prevention programs. Report submitted to The Center for Mental Health Services (SAMHSA), Prevention Research Center, Pennsylvania State University (obtainable at http://www.psu.edu/dept/prevention/)

Gresham, F. M. (2005). Response to Intervention: An Alternative Means of Identifying Students as Emotionally Disturbed. *Education & Treatment of Children*, 28(4), 328-344. Retrieved from EBSCO*host*.

Gridley, B. E., Mucha, L., & Hatfield, B. B. (1995). Best practices in preschool screening. In A. Thomas& J. Grimes (Eds.), *Best practices in school psychology–III* (pp. 213-225). Washington, DC: National Association of School Psychologists.

Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J., Richardson, R., et al. (2006). The equivalence of the Child Behavior Checklist/1 1/2-5 across parent race/ethnicity, income level, and language. *Psychological Assessment*, *18*(3), 313-323. doi:10.1037/1040-3590.18.3.313.

Gur, H. Hendin, C. P. O'Brien, M. E. P. Seligman, & B. T. Walsh (Eds.), Treating and preventing adolescent mental health disorders (pp. 579–596). New York: Oxford University Press.

Harris, J. J., Brown, E. L., & Richardson, J. W. (2004). African Americans and multicultural education: A proposed remedy for disproportionate special education placement and under inclusion in gifted education. *Education and Urban Society, 36,* 304–341. doi: 10.1177/0013124504264444

Harris-Murri, N., King, K., &Rostenberg, D. (2006). Reducing Disproportionate Minority Representation in Special Education Programs for Students with Emotional Disturbances: Toward a Culturally Responsive Response to Intervention Model. *Education &Treatment of Children*, 29(4), 779-799. Retrieved from EBSCO*host*.

Halvorsen, I., Andersen, A., & Heyerdahl, S. (2005). Girls with anorexia nervosa as young adults: Self-reported and parent-reported emotional and behavioural problems compared with siblings. *European Child & Adolescent Psychiatry*, *14*(7), 397-406. doi:10.1007/s00787-005-0489-0.

Henderson, J. (2009). Disproportionality in special education: The relationship between Prereferral Intervention Teams and the special education process(Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.(Accession Order No. AAT 3323219)

Husky, M., Kaplan, A., McGuire, L., Flynn, L., Chrostowski, C., &Olfson, M., (2011). Identifying Adolescents at Risk through Voluntary School-Based Mental Health Screening. *Journal of Adolescence*, 34(3), 505-511.

INDIVIDUALS WITH DISABILITIES EDUCATION IMPROVEMENT ACT OF 2004 (IDEA), Pub. L. 108-466.

Jaccard, J., Dittus, P., & Gordon, V. (1998). Parent-adolescent congruency in reports of adolescent sexual behavior and in communications about sexual behavior. *Child Development*, *69*(1), 247-261. Retrieved from MEDLINE with Full Text database. doi:10.1111/j.1467-8624.1998.tb06146.x

Jones, D., Dodge, K., Foster, E., Nix, R., & Conduct Problems Prevention Research Group (2002).Early identification of children at risk for costly mental health service use. *Prevention Science. 3*(4), 247-256.doi:10.1023/A:1020896607298

Jöreskog KG, Sörbom D. Chicago: Scientific Software International; 2006.

Kamphaus, R. W., Rowe, E., Dowdy, & Hendry, C.N.. (2006). Classification and Diagnosis Concepts. In R.W. Kamphaus & J.M. Campbell (Eds.). Psychodiagnostic Assessment of Children: Dimensional and Categorical Approaches. New York: John Wiley & Sons.

Kamphaus, R. W., Thorpe, J., Winsor, A. P., Kroncke, A., Dowdy, E., &VanDeventer, M. (2007). Development and predictive validity of a teacher screener for child behavioral and emotional problems at school. *Educational and Psychological Measurement*, *67*, 342–356. doi:10.1177/00131644070670021001

Kim, S., & Rowe, E. (2004, April). Similar risks, dissimilar outcomes: Rethinking risk in educational context. Paper presented at the annual conference of the American Educational Research Association, San Diego, CA.

Knight, G., Little, M., Losoya, S., &Mulvey, E. (2004). The Self-Report of Offending among Serious Juvenile Offenders: Cross-Gender, Cross-Ethnic/Race Measurement Equivalence. *Youth Violence and Juvenile Justice*, *2*(3), 273-295. doi:10.1177/1541204004265878

Lau, A., McCabe, K., Yeh, M., Garland, A., Hough, R., &Landsverk, J. (2003). Race/ethnicity and rates of self-reported maltreatment among high-risk youth in public sectors of care.*Child Maltreatment*, *8*(3), 183-194. doi:10.1177/1077559503254141.

Lam, S., & Cheng, S. (2002). Cross-informant agreement in reports of environmental behavior and the effect of cross-questioning on report accuracy. *Environment and Behavior*, *34*(4), 508-520. doi:10.1177/00116502034004006.

Lee, S., Elliott, J., & Barbour, J. (1994).A comparison of cross-informant behavior ratings in school-based diagnosis.*Behavioral Disorders*, *19*(2), 87-97. Retrieved from PsycINFO database.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review*.*10* (1)1-41.

Milan, M. (1983). English Language Independent Prediction of Academic Readiness in Ethnically Diverse Low SES Four and Five Year Old Children.

Moretti, M., Fine, S., Haley, M., Marriage, K. (1985). Childhood and Adolescent Depression: Child-report versus Parent-report Information. *Journal of the American Academy of Child Psychiatry*, *24*(3), 298-302.doi:10.1016/S0002-7138(09)61090-6.

Mustian, A.. The comparative effects of function-based versus nonfunction-based interventions on the social behavior of African American students. Ph.D. dissertation, The University of North Carolina at Charlotte, United States -- North Carolina. Retrieved March 14, 2011, from Dissertations & Theses: A&I.(Publication No. AAT 3422674).

Ollendick, T., Oswald, D., & Francis, G. (1989).Validity of teacher nominations in identifying aggressive, withdrawn, and popular children. *Journal of Clinical Child Psychology*, *18*(3), 221-229. doi:10.1207/s15374424jccp1803_4.

Pagano, M.E., Cassidy, L. J., Little, M., Myrphy, J. &Jellinek, M.S. (2000). Identifying psychosocial dysfunction in school-age children: The Pediatric Symptom Checklist as a self-report measure. Psychology In The Schools 37(2), 91-106.doi:10.1002/(SICI)1520-6807(200003)37:2<91::AIS-PTS1>3.0.CO;2-3

Pas, E. T., Bradshaw, C. P., Hershfeldt, P. A., & Leaf, P. J. (2010). A multilevel exploration of the influence of teacher efficacy and burnout on response to student problem behavior and school-based service use. *School Psychology Quarterly*, *25*(1), 13-27. doi:10.1037/a0018576

Patterson, G. R., Reid, J. B., &Dishion, T. J. (1992). Antisocial boys. Eugene, OR: Castalia Press.

Romer, D., & McIntosh, M. (2005).The roles and perspectives of school mental health professionals in promoting adolescent mental health. In D. L. Evans, E. B. Foa, R. E.

Serbin, L. A., Marchessault, K., McAffer, V., Peters, P., & Schwartzman, A. E. (1993). Patterns of social behavior on the playground in 9- to 11 year-old girls and boys: Relation to teacher perceptions and to peer raings of aggression, withdrawal, and likability. In C. H. Hart (Ed.), Children on playgrounds: Research perspectives.

Sharp, C., Croudace, T. J., Goodyer, I. M. & Amtmann, D. (2005) The Strength and Difficulties Questionnaire: Predictive validity of parent and teacher ratings for help-seeking behaviour over one year. *Educational and Child Psychology*. 22(3), 28-44.

Skiba, R., McLeskey, J., Waldron, N., Grizzle, K., & Bartley, J. (1993). The Context of failure in the primary grades: Risk factors in low and high referral rate classrooms. School Psychology Quarterly 8(2), 81-98. doi: 10.1037/h0088837

Skiba, R., Simmons, A., Ritter, S., Kohler, K., Henderson, M., & Wu, T. (2006). The Context of Minority Disproportionality: Practitioner Perspectives on Special Education Referral. *Teachers College Record*, 108(7), 1424-1459. doi:10.1111/j.1467-9620.2006.00699.x

Skotko, B. (2011, October). Prenatal and Postnatal Diagnoses of Down Syndrome: New Advances and Updates for Today's Physicians. *Children's Hospital Los Angeles Pediatric Grand Rounds*. Lecture conducted from Children's Hospital Los Angeles, Los Angeles, CA.

Sweeting, H. & West, P. (1998) Health at age 11: reports from schoolchildren and their parents. *Archives of Disease in Childhood*, *78*, 427–434.

Tejeda-Delgado, M.(2009) Teacher Efficacy, Tolerance, Gender, and Years of Experience and Special Education Referrals. *International Journal of Special Education*.*24*(1), 112-119.

Terras, M. M., Thompson L.C. & Minnis, H. (2009) Dyslexia and psycho-social functioning: An exploration study of the role of self-esteem and understanding. *Dyslexia: An International Journal of Research And Practice*, *15* (4). 304-327.doi:10.1002/dys.386

Thurlow, M. L., Sinclair M.F., Johnson, D. R., & National Center on Secondary Education and Transition, M. N. (2002). *Students with Disabilities Who Drop Out of School: Implications for Policy and Practice. Issue Brief: Examining Current Challenges in Secondary Education and Transition*

Twyford, J. M., Chin, J. K., Eklund, K. R., & Dowdy, E. (2009). Measurement Equivalence Across Ethnicities for a Teacher Behavior Rating Screener. 3. Retrieved from EBSCO*host*.

U.S. Department of Education. (2006). *26th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2004*. Jessup, MD: Ed Pubs.

Waitoller, F., Artiles, A., Cheney, D. ,Artiles, A. , & Cheney, D. (2010). The Miner's Canary: A Review of Overrepresentation Research and Explanations. *Journal of Special Education*, *44*(1), 29-49.

Walker, H. M., & Severson, H. (1990). Systematic screening for behavioral disorders (SSBD). Longmont, CO:Sopris West.

Walker, H. M., Stiller, B., Severson, H. H., Feil, E. G., & Golly, A. (1998). First step to success: Intervening at the point of school entry to prevent antisocial behavior patterns. *Psychology in the Schools*, **35**, 259–269.doi:10.1002/(SICI)1520-6807(199807)35:3<259::AID-PITS6>3.3.CO;2-B

Walker, H., Nishioka, V., Zeller, R., Severson, H., &Feil, E. (2000). Causal Factors and Potential Solutions for the Persistent Under identification of Students Having Emotional or Behavioral Disorders in the Context of Schooling. *Assessment for Effective Intervention*, *26*(1), 29-39. doi:10.1177/073724770002600105

Walker, B., Cheney, D., Stage, S., & Blum, C. (2005).Schoolwide screening and positive behavior supports: identifying and supporting students at risk for school failure. *Journal of Positive Behavior Interventions*, 7(4), 194-204. doi:10.1177/10983007050070040101

Waters, E., Stewart-Brown, S., & Fitzpatrick, R. (2003). Agreement between adolescent self-report and parent reports of health and well-being: results of an epidemiological study. *Child: Care, Health & Development*, *29*(6), 501-509. doi:10.1046/j.1365-2214.2003.00370.x

Weitkamp, K., Romer, G., Rosenthal, S., Wiegand-Grefe, S., & Daniels, J. (2010). German Screen for Child Anxiety Related Emotional Disorders (SCARED): Reliability, validity, and cross-informant agreement in a clinical sample. *Child and Adolescent Psychiatry and Mental Health*, *4*doi:10.1186/1753-2000-4-19.

vanWidenfelt, B., Goedhart, A., Treffers, P., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*(6), 281-289. doi:10.1007/s00787-003-0341-3.

CHAPTER 2
MEASUREMENT EQUIVALENCE OF THE BASC-2 BEHAVIORAL AND EMOTIONAL SCREENING SYSTEM ACROSS RACIAL GROUPS

Introduction

The Behavioral Assessment System for Children-2 Behavioral and Emotional Screening System (BASC-2 BESS) is the most recent additional to the BASC-2 family of evaluation and intervention tools. This instrument was designed to measure the risk of behavioral and emotional problems in students, yet do so in approximately 5 minutes per child. The BASC-2 BESS screening system includes Parent and Teacher forms, which are available for children in pre-school-12th grade and a Student Self-Report form that is available for students in third -12th grades. Moreover, the authors designed the instrument to be administered universally as a part of a multi-gate approach to behavioral and emotional intervention (Kamphaus & Reynolds, 2008).

Behavioral and Emotional Risk (BER) has been defined as "atypical development in comparison to children of the same age in the areas of Maladaptive behaviors, emotions, thought patterns, and delayed acquisition of pro-social and coping skills" (Kamphaus, 2011). It is important to note that BER ***does not*** meet the diagnostic criteria for a mental health disorder or the criteria for qualification for special education programming. Some examples of BER commonly observed in childhood include: mild inattention, odd/immature social skills development, atypical amounts of worry or levels of sadness, more instances of aggress than that of peers, and/or bullying. Furthermore, socio-cultural influences (e.g. poverty, unemployment, access to healthcare, language

barriers, etc) position students in racial minority groups, particularly from urban settings, for increased exposure to the risk factors that contribute to BER. By assessing BER commonly associated with the development and manifestation of behavioral and emotional disorders, the authors of the BASC-2 BESS aimed to support the use of early intervention. Early intervention has been identified as a method of deterring the pernicious outcomes of a childhood characterized by both BER and/or behavioral and emotional disorders (Blair & Diamond, 2008; Greenberg, Domitrovich, & Bumbarger, 2000).

A systematic approach to identifying students who are demonstrating BER or "universal screening" for BER is aligned with data-driven approaches for identifying students for interventions and special education services. Those students identified with BER who do not respond to empirically validated interventions in general education settings can receive more intensive educational supports in a timely fashion, thus improving their educational trajectory (Dowdy, Ritchey, & Kamphaus, 2010). The BASC-2 BESS is widely used to measure and provide an assessment of BER in children across the United States (e.g. Los Angeles Unified School District, State of New Hampshire).

Although the validity and structure of the BASC-2 BESS have been viewed favorably in the literature (Distefano & Morgan, 2010; Dowdy et al, 2011; Twyford, Chin, Eklund, & Dowdy, 2009), the present body of research has very few peer-reviewed studies that examine the BASC-2 BESS across population subgroups. Specifically, measurement equivalence studies across groups (e.g. gender, race, culture, language) have been scarcely addressed in the literature. Measurement equivalence studies explore

the nature of an instrument to assess the utility of the instrument with various groups (e.g. gender, race, etc.). Hu and Triandis (1985) summarize cross-cultural measurement equivalence into four types of equivalence commonly sought after in psychology these are: *1) Conceptual/functional equivalence 2)Equivalence in construct operationalization 3) Item equivalence 4) Scalar equivalence*. Conceptual/functional measurement equivalence evaluates the degree to which the construct measured by the instrument is perceived and functions across different groups. This type of equivalence assesses the antecedents and consequences for a construct as well as the presence of a "universal learning situation" and goals across groups. This type of equivalence aims to determine if each construct holds the same meaning for each group and may do so by observing the relationship between latent and observed variables. Conceptual/functional equivalence also may also assess circumstances surrounding how scores are derived often by comparing instruments to others seeking to measure the same construct. Equivalence in construct operationalization bridges the gap from theory to measurement by assessing if a construct can, in essence, be generalized (e.g. same operational definition and value) across cultures. For example, operationalizing the construct of "somatization" as the verbal report of physical discomfort would lack equivalence if the aim were to compare somatic complaints of nonverbal and verbal populations. Item equivalence assumes both construct/functional equivalence and operationalization have been established. This more specific type of measurement equivalence explores the meaning of the individual items across groups to assess the validity of the scores that are derived. Finally, scalar measurement equivalence, the most difficult to truly achieve, assumes the aforementioned equivalence have been achieved and seeks to measure the degree or level

of intensity the construct resonates across each group. Researchers may use regression analysis to explore linear relationships between instruments to explore this type of equivalence (Hui & Triandis, 1985; Jastrowski-Mano, Hobert-Davies, Klein-Tasman, & Adesso, 2009). Establishing these various levels of equivalence is of great importance to screening for BER. A lack of measurement equivalence would have great implications for the use of such instruments with specific groups. If measurement equivalence in an instrument is not found and constructs are manifesting differently across groups, the use of the instrument should be scrutinized in those groups.

Dowdy and colleagues (2011) revealed through Exploratory Factor Analysis (EFA) that the items on this BESS Student measure are best aligned with a four factor model including, “Internalizing Problems”, “Personal Adjustment”, “School Problems”, and “Inattention/Hyperactivity.” Each of these latent factors is aligned with domains on the BASC-2. Dowdy and colleagues study also employed Confirmatory Factor Analysis (CFA) to assess the validity of the EFA using a portion of the BESS Student norming sample (sample used by publishers to establish the initial reliability and validity of the instrument), and a sample taken from a large urban school district. These findings imply the BESS Student items load onto the same factors in both samples and found the instrument to assess the same constructs across the two groups. . Consequently, it appears that there is pressing need for more published evidence on the measurement equivalence of the BESS Student in regards to diverse groups.

While there are a number of statistical strategies that can be employed to assess measurement equivalence for the purposes of this research, this researcher seeks to work toward construct/functional measurement equivalence using Confirmatory Factor

Analysis (CFA). Specifically, employing multi-group CFA that tests the assumption of invariance of the four-factor model across all groups assessed. In an in depth query of the structure of the BASC -2 BESS student form,

Measurement equivalence across racial groups is foundational for supporting the score inferences of test. Many studies have assessed student behaviors or informant report information by student race (Adams, Kelley, & McCarthy. 1997; Knight, Little, Losoya, & Mulvey, 2004;Lau, McCabe, Yeh, Garland, Hough, & Landsverk, 2003). The outcomes of measurement equivalence studies of parent report of behavioral functioning have repeatedly revealed differences by racial group (Achenbach, McConaughy, & Howell, 1987; Gross, Fogg, Young, Ridge, Cowell, & Richardson et al, 2006). Achenbach (2000) reported elevated scores on 15 items of the CBCL in parent reports for racial minority pre-school children. It has been difficult to determine if racial differences in parent reporting are a result of variance in interpretation of questions based on ethnic groups or other factors. Nevertheless, the information gathered is vital for psychological practice. Evidence of measurement nonequivalence by informant has been found to be particularly evident in instruments that use the same cutoff score for all participants and do not provide different cut offs for different groups (Gross, Fogg, Young, Ridge, Cowell, & Richardson et al, 2006).

This study specifically aims to establish measurement equivalence of the BESS Student form across multiple racial groups. The use of the BESS Student form was selected for several reasons. First, self-report is frequently used as a tool for gathering information about behavioral and emotional functioning and is considered to yield accurate information regarding self-perception of behavioral and emotional functioning

(Adams, Kelley, & McCarthy. 1997; Knight, Little, Losoya, & Mulvey, 2004)). Second, self-report has been found to be an optimal method for gathering information about students with internalizing symptoms (Mays, 2008; Merrell, McClun, Kempf and Lund, 2002) Third, self-report surveys or questionnaires reduce the potential for interviewer bias and in most cases limit the influence of social desirability in participant responses. Interviewer bias and social desirability commonly interfere with results when information is gathered through in-person interviewers (Blount, Evans, Birch, Warren, & Norton, 2002). Fourth, self-report surveys or questionnaires like the BESS are generally low cost and easily distributed to a large number of subjects (i.e. practical). Furthermore, scoring technology supports the use of quick scoring and interpretation of results versus the time need to score qualitative interview methods. Fifth, self- reports have been found to predict student achievement (Carroll et al, 2009; Martin and Debus, 1998). Finally, self-reports are a valuable and preferred method of gathering information for personality and behavioral data (Achenbach, 2006; Blount, 2002).

Additionally, in the context of universal screening, evidence supporting the validity of self-report forms would suggest the feasible use of universal screening for gathering mental health risk information. In schools, gathering information about all students from teachers and parents may be daunting and impractical from both time and fiscal perspectives. Also, obtaining information about student mental health and other sensitive information can be precarious when using parents or teachers as informants. In a study of parent awareness of suicidal ideation, Mojtabi and Olfson (2008) found that roughly 60% of parents were unaware of their child's suicidal ideations. In other studies, parental reporting of behavioral and emotional functioning has been found to vary from

self-report results (Moretti, 1985; Waters, Stewart-Brown, & Fitzpatrick, 2003). Similar to parents, research supporting teachers as informants is unconvincing. Researchers have demonstrated that teachers are inconsistent as informants (Eklund, et al., 2009; Epkins & Meyers 1994). On the other hand, teacher reports of behavior have been found to be more predictive of behavioral and emotional outcomes than parent reports, but are not consistently aligned with self-report (Sharp, Croudace, Goodyer, &Amtmann, 2005). Ollendick, Oswald, and Francis (1989) found in their assessment of risk for behavioral and emotional problems that teacher report yielded similar findings to peer and self-report of behavior and risk. However, when compared to self-report in a study of depressive symptoms in Chinese children, Tepper, Guo, Zhai, Liu, & Li (2008), found teacher report to endorse substantially fewer symptoms.

Students, particularly adolescents, tend to be more forthright when reporting with paper and pencil than in interview (Husky et al., 2011). Youth are also generally unlikely to spontaneously disclose symptoms associated with BER to parents and teachers (Husky et al, 2011; Waters, Stewart-Brown, & Fitzpatrick, 2003). Gathering information directly from students offers direct knowledge of their perceptions of their behavioral and emotional functioning and can be instrumental in designing interventions to promote their growth and success. For these reasons, among others, exploring the measurement equivalence of self-report forms is pivotal to the investigation of BER screening and the impact of BER on student performance.

In general, there is a pressing need for research on the measurement equivalence of universal screening tools across racial groups. Measurement equivalence studies have been conducted on an array of instruments designed to diagnose and assess the severity of

behavioral and emotional disorders such as the Child Behavior Checklist or the Beck Depression Inventory (Jastrowski-Mano, Hobert-Davies, Klein-Tasman, &Adesso, 2009; Stapleton, Sander, & Stark, 2007). However, minimal research is available on the psychometric properties of the instruments designed to measure BER. The purpose of this paper is to add to the empirical literature of the BESS by examining its factor structure in three different samples of school-age students using the BESS Student form. In doing so, this paper uses theory-driven Confirmatory Factor Analysis (CFA) models that map onto previously identified BESS factors, and cross-validates these models across three samples using CFA. Thus, the research questions are:

1) Does the four -factor model of the BESS Student emerge in norming sample using only the Black, Hispanic, and White participants?

2) Do the BESS Student items load onto the same factors for each individual group (Black, White, and Hispanic)?

The results from this study will be useful for helping psychological practitioners and school based referral team members, especially school psychologists, determine the validity of inferences based on Self-Report BESS scores for diverse groups of students in their own practice. Additionally, the results may further support the use of a universal screening tool such as the BESS for norm-referenced identification of students at risk for behavioral and emotional problems across the three predominant racial/ethnic groups currently in the US.

Methods

Sample

The sample for this study was taken from the larger BESS Student Form norming data set, which was composed of students ages 8–18 and is representative of the population of the United States. The sample included racial categories reported by parents as follows: 13% Black (*N* =714), 18% Hispanic (N = 1,025), and 69% White (*N* = 3,939). These three groups were utilized for multi-group CFA to assess the structure for each group. Only students of these three racial groups were selected for this present study due to small representation of other groups (e.g. Asian, Multi-Racial, Other) in the sample. The total gender representation of the three groups was 47% male participants (total *N* = 5,678). Gender representation was approximately equal across all races and ages (see Table 1).

Table 1

Demographics from Norm Sample race and gender

Race	Total *N*	Female *n*
Black	714	54%
Hispanic	1126	54%
White	3838	52%

Measure

The BESS Student Form (Kamphaus & Reynolds, 2007; BESS Student) is a 30-item instrument designed to measure self-reported levels of risk for behavioral and

emotional problems. The BESS Student requires no informant training, can be completed less than 10minutes, and is available in both Spanish and English. The BESS was developed using a norming sample of 12,350 teacher, parent, and student forms, collected from 233 cities in 40 states. The BESS Student norming sample includes students in Grades 3 through 12 reporting on their behavioral and emotional functioning using a 4-point scale (i.e., *never, sometimes, often, almost always*). Summing the responses to the problem items and the reverse scores of the adaptive behavior items creates a raw score. The raw score is transformed to a total *T*-score, in which higher scores reflect more problems; 20-60 suggests a "Normal" level of risk, 61-70 suggests an "Elevated" level of risk, and scores of 71 or higher suggests an "Extremely Elevated" level of risk. These classification labels of risk were determined according to a normal distribution of the norming sample scores and the distance of the scores from the normative sample mean. This method of classification was developed with the intention of assisting practitioners with decision-making regarding students who may require additional assessment and intervention (Kamphaus & Reynolds, 2007). Although Parent and Teacher Forms are also available, this study focused solely on the BESS Student.

The psychometric properties of the BESS Student are generally acceptable, having good split-half reliability (.90-.93) and test-retest reliability (.80). The BESS Student has moderate correlations with total scores from other measures of behavioral and emotional problems, including the Achenbach System of Empirically Based Assessment (ASEBA) Youth Self Report Form (.81). The test manual also reports classification accuracy when using the BESS Student to predict full BASC-2 Self Report of Personality problem composites as having moderate sensitivity, high specificity,

moderate positive predictive value (PPV), and high negative predictive value (NPV). The authors also report that the sensitivity, specificity, PPV and NPV tend to be highest when predicting the Emotional Symptoms Index (ESI) and Internalizing Problems from the BASC-2 SRP. For additional information regarding the psychometric properties of the BESS, readers are referred to the BESS Manual (Kamphaus & Reynolds, 2007). Dowdy (2011) and colleagues provided evidence that the BESS Student has a four-factor structure through EFA and CFA.

Procedure

The BESS Student Form was normed on a large sample that is representative of the general population of U.S. children with regard to sex, race/ethnicity, and clinical or special education classification (Kamphaus & Reynolds, 2007). Using multi-group CFA, this study explored the presence of these factors in three racial groups taken from the norming sample. For the purposes of this study, all students who were reported to be Black, Hispanic, and White were selected from the norming sample. The first 50% of the dataset (*N*= 2,839) was utilized in the development of the model and the second half (*N*= 2,839) of the dataset was used to confirm the model.

Data Analysis Plan

A series of Confirmatory Factor Analyses were computed comparing the factor structure indicated by Dowdy and colleagues (2011) with other plausible models. In the initial model, variables specified for each factor were permitted to freely correlate, with the exception of the reference variable for each factor, which was set to 1.0. The item with the highest loading variable was identified and used as a reference for other parameters. The goal was to determine if the factor pattern, factor loadings, factor

correlations, and correlated errors were equivalent in Black, Hispanic, and White samples.

Mplus software (version 5.21, Muthén & Muthén, 2004) was used to perform the multi-group CFA. Maximum Likelihood estimation was used because it is generally robust providing goodness-of-fit indices, weighted mean and variance adjustment for factor loadings, and correlations (Finney & DiStefano, 2006). Additionally, this estimation was utilized in the Dowdy et al, study. Figure 2illustrates Dowdy and colleagues' final four-factor model (Personal Adjustment, Inattention/Hyperactivity, Internalizing Problems, and School Problems) tested with the multi-group CFA. Based on Dowdy's model, each factor has varying number of corresponding items. The four factors and descriptions of corresponding items are available in Table 2. Additionally, the Dowdy et al model omitted: Item 9 (Being liked by others), Item 11(Difficulty sitting still), and Item 22(Feeling stupid). These items were deemed problematic as they yielded factor loadings greater than 1.0.

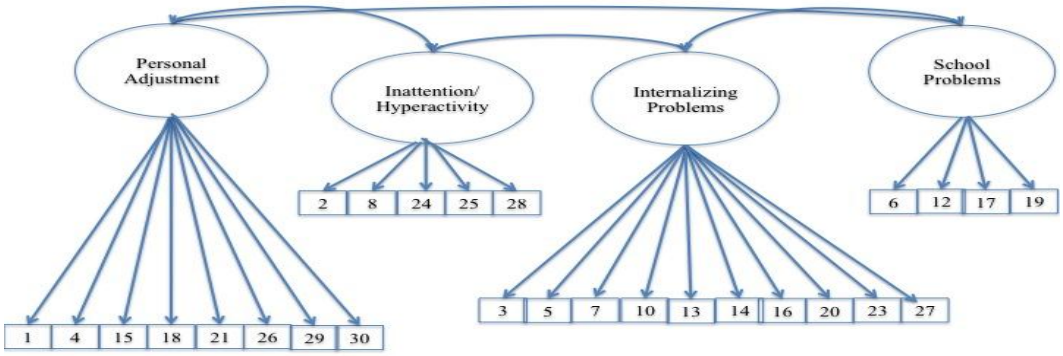

Figure 1. Path diagram of final Exploratory Factor Analysis Model (Dowdy et al, 2011)

Table 2

BESS Student Items and Factor - Dowdy et al (2011) CFA

Item #	Item Description	Factor
1	Good at decision making	Personal Adjustment
4	Like looks	Personal Adjustment
15	Parental trust	Personal Adjustment
18	Parents listen	Personal Adjustment
21	Others think I'm fun to be with	Personal Adjustment
26	Parents are proud	Personal Adjustment
29	School comfort	Personal Adjustment
30	Others respect me	Personal Adjustment
2	Talk when others talk	Inattention/Hyperactivity
8	Paying attention to teacher	Inattention/Hyperactivity
24	Noisy	Inattention/Hyperactivity
25	Trouble for inattention	Inattention/Hyperactivity
28	Difficulty standing still	Inattention/Hyperactivity
3	Worries	Internalizing Problems
5	Feeling out of place	Internalizing Problems
7	Others angry at	Internalizing Problems

10	Life getting worse	Internalizing Problems
13	People out to get me	Internalizing Problems
14	Worry about future	Internalizing Problems
16	Feeling left out	Internalizing Problems
20	Wanting to improve, but unsuccessful	Internalizing Problems
23	Blamed for problems out of my control	Internalizing Problems
27	Failure despite effort	Internalizing Problems
6	Interest in quitting school	School Problems
12	School interest	School Problems
17	Hate school	School Problems
19	Unfair teachers	School Problems

The present study used multiple indices to evaluate model fit based upon recommendations in the literature (Hu & Bentler, 1995), Comparative Fit Index (CFI), the standardized root mean square residual (SRMR), the root mean square error of approximation (RMSEA) and Chi-square (X^2). The CFI provides a measure of how well the hypothesized structure corresponds with the model and avoids underestimation of fit. Values greater than .90 for the CFI are typically required to indicate a good fit (Jöreskog & Sörbom, 2006). The SRMR represents the standardized discrepancy between the observed covariance and the predicted covariance matrices. Values of .05 or less suggest good model fit. RMSEA provides an additional model fit index relative to the population

covariance matrix accounting for the complexity of the model. Values less than .05 for the RMSEA indicate good fit, with values as high as .08 representing a reasonable fit. Finally the X^2statistic is used to measure the level of significance of fit between each sequential model (Hu & Bentler, 1999; Jöreskog & Sörbom, 1993).

Additionally, Multivariate Analyses of Variance (MANOVA) was used to determine the existence of statistically significant mean differences on the BESS Student *T*-scores and the student race/ethnicity.

Results

Model 1. The first model tested is depicted in Figure 3. In this model, all items were set to load onto their respective factors as derived from the larger BASC-2 factor structure. Although Dowdy et al. (2011) allowed for correlated errors in their final model, Model 1 was tested as the base hypothesis in this study in order to determine if these modifications were appropriate for the present sample. In Model 1, all paths and relationships were constrained to be equal across the three ethnic/racial groups. Although the chi–square test of model fit was significant χ^2 (1369)=6223.132, other fit indices were used to assess model fit due to the large sample size (e.g. Jöreskog, 1993). For Model 1, the fit indices suggested an unacceptable fit, CFI = .810, RMSEA = .061, SRMR = .066. Because the fit indices failed to meet the criteria for a good fit, modification indices were considered to improve model fit. Modification indices suggested that Item 9 was problematic, as Dowdy and colleagues indicated.

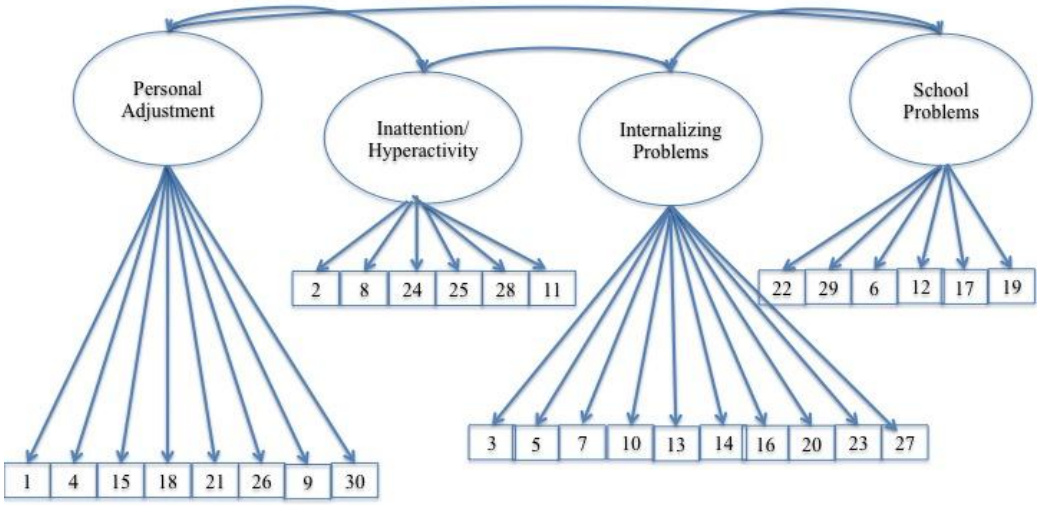

Figure 2. Model 1

Model 2. In Model 2, item 9 was omitted both to be consistent with the results of previous CFA analyses of the BESS Student (Dowdy et al, 2011) and due to the empirical results of Model 1. This item deletion was the only change from Model 1 to Model 2; all else remained consistent. Similarly, to the fit of Model 1 the chi–square test of model fit was significant in Model 2 χ^2 (1279)= 5085.890. However, the fit indices suggested that the model was still unacceptable despite the modification, CFI = .843, RMSEA = .056, SRMR = .060. Because the fit indices failed to meet the criteria for a good fit, modification indices were considered to improve model fit. Model 2 was statistically

superior to Model 1 ($\Delta\chi^2$= 1137.242; *df* = 90; see Table 3). Modification indices suggested that item 11 was indeed problematic, as Dowdy et al suggested.

Model 3. In Model 3, both items 9 and 11 were omitted in accordance with the previous CFA analysis of the BESS Student. This second deletion was the only change from Model 2 to Model 3; all else remained consistent. Again, the chi–square test of model fit was found to be significant χ^2 (1192) = 4699.282. The other fit indices, however, suggested an unacceptable fit, CFI = .847, RMSEA = .056, SRMR = .059. Therefore, modification indices were considered to improve model fit. Model 3 was statistically superior to Model 2($\Delta\chi^2$(87)= 386.608;see Table 3). Modification indices suggested that item 22 was also problematic, similar to the findings of Dowdy and colleagues (2011).

Model 4. In Model 4, items 9, 11, and 22 were simultaneously omitted in accordance with the previous CFA analysis of the BESS Student. This additional deletion was the only change from Model 3 to Model 4; all else remained consistent. The chi–square test of model fit for Model 4 was significant χ^2 (1108)= 4183.095. Yet upon review of the fit indices the model fit for Model 4 were also found to be unacceptable, CFI = .859, RMSEA = .054, SRMR = .056 and modification indices were suggested to improve the fit. Model 4 was statistically superior to Model 3 ($\Delta\chi^2$= 516.187; df = 84; see Table 2). Modification indices suggested item 10 was cross loading onto Factor 1 (Personal Adjustment) for the Hispanic group in addition to Factor 3 (Internalizing Problems) as anticipated.

Model 5. In Model 5, items 9,11, and 22 were omitted in accordance with the previous CFA analysis of the BESS Student. Additionally, item 10 was permitted to

cross-load onto Factor 1 in the Hispanic group. This cross loading was the only change from Model 4 to Model 5; all else remained consistent. The Model 5 chi–square test of model fit was again significant, χ^2 (1107)= 4162.807. The fit indices of Model 5 revealed that fit was also unacceptable, CFI = .860, RMSEA = .054, SRMR = .056. Model 5 was statistically superior to Model 4 ($\Delta\chi^2$= 20.288;*df*= 1; see Table 2). Modification indices suggested that item 29 was also cross loading onto Factor 1 (Personal Adjustment) in the Hispanic group as well as Factor 3 (Internalizing problems).

Model 6. In Model 6, items 9,11, and 22 were omitted in accordance with the previous CFA analysis of the BESS Student. Additionally, items 10 and 29 were permitted to cross-load onto Factor 1 in the Hispanic group. This additional cross loading was the only change from Model 5 to Model 6; all else remained consistent. The chi–square test of model fit for Model 6 was statistically significant, χ^2 (1106) = 4124.210. Again, this model was found to be unacceptable upon review of the fit indices, CFI = .862, RMSEA = .054, SRMR = .056. Model 6 was statistically superior to Model 5 ($\Delta\chi^2$= 37.79;*df* = 1; see Table 2). Modification indices suggested that item 10 was cross loading onto Factor 1 (Personal Adjustment) in the White group as well.

Model 7. In Model 7, items 9,11, and 22 were omitted in accordance with the previous CFA analysis of the BESS Student. Items 10 and 29 were permitted to cross-load onto Factor 1 in the Hispanic group. Item 10 was also permitted to cross-load onto Factor 1 in the White group. The additional cross loading in the White group was the only change from Model 6 to Model 7; all else remained consistent. The chi–square test of model fit of Model 7 was significant, χ^2 (1105)= 4072.049, and the remaining fit indices suggested an unacceptable fit, CFI = .864, RMSEA = .053, SRMR = .055. Model

7 was statistically superior to Model 6 ($\Delta\chi^2$= 52.161; *df* = 1; see Table 3). Modification indices suggested that item 29 was also cross-loading onto Factor 1-(Personal Adjustment) in the White group.

Model 8. In Model 8, items 9,11, and 22 were omitted in accordance with the previous CFA analysis of the BESS Student. Items 10 and 29 were permitted to cross-load onto Factor 1 in the Hispanic group and the White group. The additional cross loading of item 29 in the White group was the only change from Model 7 to Model 8; all else remained consistent. The fit of Model 8 yielded a significant chi-square, χ^2 (1104)= 3967.917. The remaining fit indices suggested unacceptable fit, CFI = .869, RMSEA = .052, SRMR = .055. Model 8 was statistically superior to Model 7 ($\Delta\chi^2$= 104.573; *df* = 1; see Table 3). Modification indices suggested that item 6 was loading onto Factor 3 (Internalizing Problems) in the White group in addition to Factor 4 (School Problems) as predicted.

Model 9. In Model 9, items 9,11, and 22 were omitted in accordance with the previous CFA analysis of the BESS Student. Items 10 and 29 were permitted to cross-load onto Factor 1 in the Hispanic group and the White group. The additional cross loading of item 6 onto Factor 3 in the White group was the only change from Model 8 to Model 9; all else remained consistent. The fit of Model 9 had a statistically significant chi-square, χ^2 (1103)= 3849.086. Yet, while they moved closer to the desired fit may be considered acceptable, for the purposes of this research, a model with a better fit was sought. Model 9 yielded fit indices: CFI = .874, RMSEA = .051, SRMR = .053. Model 9 was statistically superior to Model 8 ($\Delta\chi^2$= 118.831; *df* = 1; see Table 3). Modification

indices suggested that residual variances (items 21 and 30) should be allowed to correlate in the White group, consistent with Dowdy and colleagues' (2011) overall model.

Model 10. In Model 10, items 9,11, and 22 were omitted in accordance with the previous CFA analysis of the BESS Student. Items 10 and 29 were permitted to cross-load onto Factor 1 in the Hispanic group (Figure 4) and the White group. This model also allowed the cross loading of item 6 onto Factor 3 in the White group (Figure 5). No additional cross loadings were necessary for the Black group (Figure 6). Permitting correlated residual variances (items 21 and 30) was the only change from Model 9 to Model 10; all else remained consistent. The fit of Model 10 was unsatisfactory in terms of a significant chi-square, χ^2 (1102)= 3967.917. The fit indices were closer approximations to the desired values, CFI = .874, RMSEA = .051, SRMR = .053. Model 10 was statistically superior to Model 9 ($\Delta\chi^2$= 160.204;*df* = 1; see Table 3). This model was also statistically superior to the findings of EFA completed by Dowdy and colleagues who determined the four factor model was best support using the following Goodness-of-Fit indices: x^2 (249) = 528.705; *p* =. 00; RMSEA = .038 (90% CI =. 33–. 042), and SRMR = .028. Furthermore, Dowdy completed a CFA on a randomly selected portion of the norming sample. The model derived from the CFA yielded Goodness-of-Fit indices that were unsatisfactory according to the parameters set for this study: x^2 (316) = 644.53; CFI - .945, RMSEA = .031 (90% CI = .027–. 034); SRMR = .038. This would suggest that the specifications used for model 10 are superior for measuring the equivalence of the BESS Student across groups.

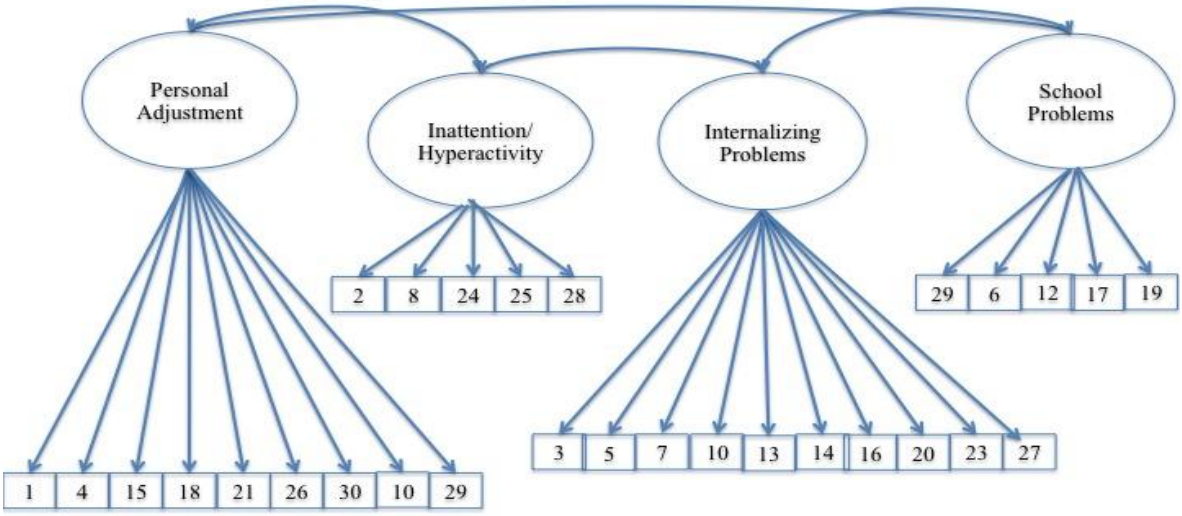

*Figure 3.*CFA Model 10 for students identified as Hispanic for the national norming sample of the BASC-2 BESS

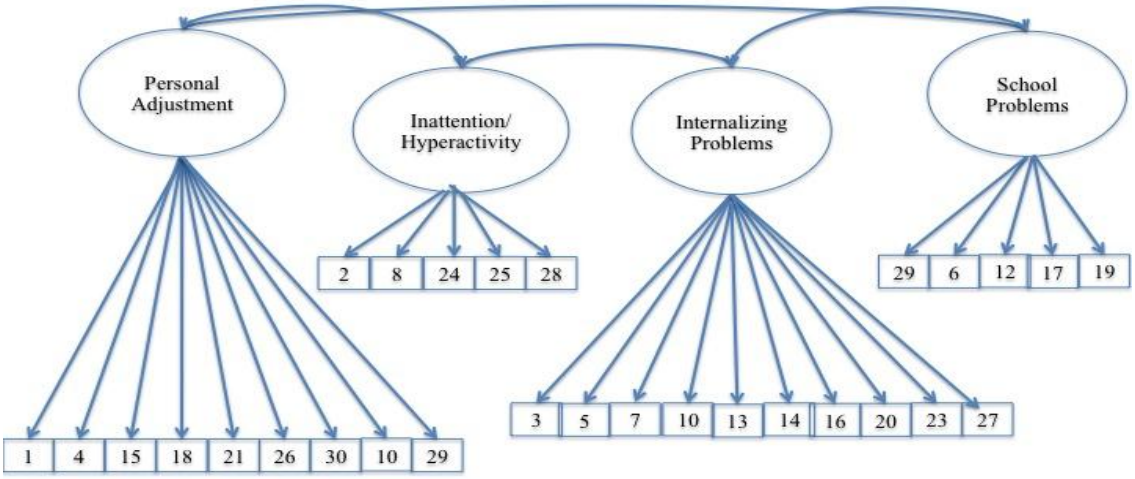

*Figure 4.*CFA Model 10 for students identified as White for the national norming sample of the BASC-2 BESS.

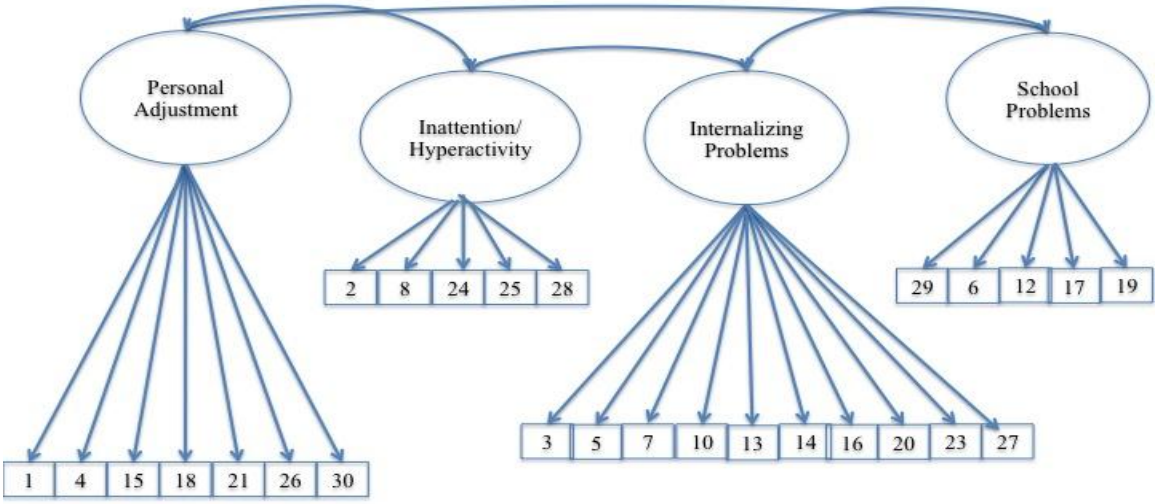

*Figure 5.*CFA Model 10 for students identified as Black for the national norming sample of the BASC-2 BESS.

Table 3
Model Fit Indices for Confirmatory Factor Analysis Models

Models	X^2 *(df)*	CFI	SRMR	RMSEA (CI 90%)	X^2difference[a]
Model 1	6223.132 (1369)	.81	.066	.061 (.060-063)	---
Model 2	5085.890 (1279)	.843	.060	.056 (.054-.058)	1137.242
Model 3	4699.282 (1192)	.847	.059	.056 (.054-.057)	386.608

Model 4	4183.095	.859	.056	.054	516.187
	(1108)			(.052-.056)	
Model 5	4162.807	.860	.056	.054	20.288
	(1107)			(.052-.056)	
Model 6	4124.210	.862	.056	.054	37.790
	(1106)			(.052-.056)	
Model 7	4072.049	.864	.055	.053	52.161
	(1105)			(.052-.055)	
Model 8	3967.917	.869	.055	.052	104.573
	(1104)			(.051-.054)	
Model 9	3849.086	.874	.053	.051	118.831
	(1103)			(.050-.053)	
Model 10	3688.882	.882	.053	.050	160.204
	(1102)			(.048-.052)	

Note. X^2= Chi-square test of model fit; *df* = Degrees of Freedom ; SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; CI =confidence interval at 90%; *p* = <.001.[a] Difference is between corresponding model and previous model.

The means and standard deviations of BESS Student *T*-scores for each factor by student race and significant interactions are presented in Table 4. In addition, the multivariate test suggested student race was not a significant predictor of mean Total BESS Student *T*-score. MANOVA analyses revealed significant differences in the mean scores for white

students compared to Hispanic students on Personal Adjustment and School Problem BESS mean scores. When compared to Black students, no significant differences were observed. Pair-wise comparisons were considered for the racial groups to determine where significant differences were present. For the BESS Student, White students reported significantly different *T*-scores than their Hispanic peers ($p = .001$) in the area of Personal Adjustment. Additionally, White students also scored significantly different than their Hispanic peers in the area of School Problems ($p = .008$).

Table 4
MANOVA Descriptive Statistics and Significance

BESS Student Score	Race	*Mean*	*SD*	*n*	*p*
Total BESS Score	Black	0.8783	0.436	360	1.000
	Hispanic	0.8830	0.419	567	1.000
	White	0.8777	0.441	1908	
	Total	0.8789	0.436	2835	
Personal Adjustment	Black	0.9709	0.566	360	1.000
	Hispanic	1.0446	0.591	567	.001*
	White	0.9483	0.571	1908	
	Total	0.9704	0.576	2835	
Inattention/Hyperactivity	Black	0.9051	0.560	360	0.918
	Hispanic	0.8340	0.561	567	0.439
	White	0.8725	0.550	1908	
	Total	0.8690	0.554	2835	

School Problems	Black	0.8082	0.645	360	0.648
	Hispanic	0.7878	0.691	567	.008*
	White	0.7891	0.684	1908	
	Total	0.7913	0.681	2835	
Internalizing	Black	0.8783	0.528	360	1.000
	Hispanic	0.8830	0.510	567	1.000
	White	0.8777	0.518	1908	
	Total	0.8789	0.518	2835	

Note. *p* values reported represent the interaction with the White group. * = significant difference

Discussion

The premise proposed at the outset of this work asserted that replacing narrative or anecdotal teacher referral practices with data-driven, norm-referenced identification methods such as universal screening holds the potential to diminish the overrepresentation of minorities in special education programs for emotional and behavior disorders. The present study sought to provide evidence regarding the measurement equivalence of the BASC-2 BESS Student form across three racial/ethnic groups: Black, Hispanic, and White. It is imperative to establish measurement equivalence to investigate the appropriateness of use of this instrument across groups. Thorough evaluation of universal screening instruments and their internal properties is

imperative if they are to be used as a tool for promoting optimal mental health in students. The exploration of the factors that emerge from the BESS Student Form across each of the racial groups may provide additional support for this instrument as a measure of behavioral and emotional risk, albeit with certain limitations and implications for future research

A previously conducted Exploratory Factor Analysis and Confirmatory Factor Analysis revealed a four-factor solution (i.e., Personal Adjustment, Inattention/Hyperactivity, Internalizing Problems, and School Problems) overall for the BESS Student (Dowdy et al, 2011). This study reproduced those findings separately for the Black, Hispanic, and White members of the norming sample by conducting a multi-group CFA across each of the racial groups with some minor allowances for each group. Unlike the Dowdy study, the model established by this study did not correlate the residual variances of items 12 (school interest) and 29 (school comfort), as this was not indicated as a statistically significant modification for any subsample in the study population.

Overall, the results indicate that the BESS Student appears to be measuring comparable constructs of risk across each of the three racial groups. While the MANOVA indicates mean differences in the White and Hispanic groups on specific factors, it would appear, based on the Total mean score, that the construct of BER is being captured in each group. Additionally, the underlying factors of Personal Adjustment, Inattention/Hyperactivity, Internalizing Problems, and School Problems as related to behavioral and emotional risk are captured in the BESS Student for each population subgroup. This finding suggests that in a brief (less than 10 min) self-rating

scale format, practitioners may be able to identify students who possess risk for a wide range of internalizing, externalizing, and school problems to guide and support the development of interventions for students in these three racial groups.

However, within each group, there are slight variations in the structure of BER. Specifically, in the Hispanic and White groups, items designed to assess risk for Personal Adjustment appeared to also assess risk for Internalizing Problems and vice versa. Item 10 (my life is getting worse) and Item 29 (school comfort) In the White group, Item 6 (Interest in quitting school) aimed at assessing risk for School Problems also appeared to assess Internalizing Problems. These findings are not completely unexpected as School Problems (academic and relational difficulties) and Personal Adjustment (self-efficacy and self-esteem) difficulties may be comorbid with Internalizing problems (Ackerman, Izard, Kobek, Brown, & Smith, 2007; Kuperminc, Leadbeater, & Blatt, 2001). These findings support the need for additional research investigating the nuances of Internalizing Problems as expressed by different subgroups and the influence of this construct on other areas of BER. Furthermore, replication of these results is important to provide further evidence that these cross-loadings are consistent across samples. The results of the present study suggest that the BESS items are loading as anticipated for students in the Black group.

Despite the generally favorable findings supporting the use of this measure across groups, further research is needed. The use of statistical methods such as Differential Item Functioning (DIF) as conceptualized in Item-Response Theory (IRT) is warranted to further explore the constructs measured by the BESS in different racial groups (Hui & Triandis, 1985). This method of analysis will inform the inner workings of each item for

each group of students providing more specific measurement equivalence. IRT results can also contribute information that can be used to increase reliability of the instrument across racial groups. Finally, IRT, when used in conjunction with additional factor analytic studies, may provide insight into the unexpected additional cross-loadings of specific items in different groups (e.g., additional loading of items 10 and 29 on the Personal Adjustment Factor for Hispanic students).

The Dowdy et al (2011) study was able to replicate the four-factor factor structure across two ethnically and regionally diverse samples using CFA methods providing preliminary validity evidence for the cross-cultural use of the BESS Student Form. This study was able to reproduce their CFA findings using the Black, Hispanic, and White participants from the norming sample. This study further strengthens the evidence for use of the BASC-2 BESS across the three largest cultural subgroups residing in the United States. To further reinforce this instrument as a sound measure across cultural groups, studies investigating the measurement invariance, different item functioning, and multiple group factor analyses including the groups that were less represented in the norming sample (e.g. Asian, Native American, Multi-racial) should be considered.

Future research should also explore the latent factor structure of the BESS Teacher and Parent forms with consideration for the race of the informant as well as the child or adolescent rated. Many studies have assessed student behaviors or informant report information by student race (Adams, Kelley, & McCarthy, 1997; Knight, Little, Losoya, & Mulvey, 2004; Lau, McCabe, Yeh, Garland, Hough, & Landsverk, 2003). Parent reports of behavior consistently show differences in behavioral rating scale outcomes by racial group (Achenbach, McConaughy, & Howell, 1987; Gross, Fogg,

Young, Ridge, Cowell, & Richardson et al, 2006). Determining if the outcomes of the BESS Parent and Teacher Forms are influenced by the race of the informant is of great consequence and has vital implications for the use of these forms in universal screening.

Mays (2008) pointed out that the body of research on the predictive validity of screening instruments like the BESS Student is limited. Continuing research on the predictive power of this instrument and other universal screening instruments and their internal properties is imperative (Dowdy et al, 2011; Glover & Albers, 2007). The information obtained from such research contributes to the understanding of the interactions between emergent factors and BER. This knowledge can also be used to advance the development and use of identification practices, assessment and early interventions for students with BER.

The findings of the CFA suggest that the BESS Student is, as designed, identifying behavioral and emotional risk across each of the three groups explored. This study is encouraging as it reveals that the use of the BESS Student as a universal screening measure could lessen inconsistencies in present special education referral practices. The BASC-2 BESS as a universal screening tool is also poised alleviate the disproportionate number of children of color identified by schools as having behavior and emotional disorders. This instrument also provides an avenue to identify students with internalizing disorders who are often overlooked in present referral practices (Bradshaw, Buckley, & Ialongo, 2008; Kataoka, Zhang, & Wells, 2002). These findings support the use of a universal screening measure as the first step in a multi-step identification and intervention process. Following up with additional assessment to evaluate the specific

areas of risk warranting intervention is pivotal to providing appropriate support services and promoting the behavioral and emotional health of students.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232. doi:10.1037/0033-2909.101.2.213

Ackerman, B. P., Izard, C. E., Kobak, R., Brown, E. D., & Smith, C. (2007). Relation between Reading Problems and Internalizing Behavior in School for Preadolescent Children from Economically Disadvantaged Families. *Child Development*, *78*(2), 581-596.

Adams, C., Kelley, M., & McCarthy, M. (1997). The Adolescent Behavior Checklist: Development and initial psychometric properties of a self-report measure for adolescents with ADHD.*Journal of Clinical Child Psychology*, *26*(1), 77-86. doi:10.1207/s15374424jccp2601_8

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology*, *20*(3), 899-911

Bradshaw, C. P., Buckley, J. A., & Ialongo, N. S. (2008). School-based service utilization among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly*, *23*(2), 169-186. doi:10.1037/1045-3830.23.2.169

Bryne, B., & van de Vijver, F. (2010). Testing for Measurement and Structural Equivalence in Large-Scale Cross-Cultural Studies: Addressing the Issue of Nonequivalence. *International Journal of Testing*,*10*(2), 107-132. doi:1.180/15305051003637306

Carroll, A., Houghton, S., Wood, R., Unsworth, K., Hattie, J., Gordon, L., & Bower, J. (2009). Self-Efficacy and Academic Achievement in Australian High School Students: The Mediating Effects of Academic Aspirations and Delinquency. *Journal of Adolescence*, *32*(4), 797-817.

DiStefano, C. & Morgan, G. (2010). Evaluation of the BESS TRS-CA Using Rasch Rating Scale Model. *School Psychology Quarterly*, *25* (4), 202-212.doi:10.1037/a0021509.

DiStefano, C. & Morgan, G. (2011). Examining Classification Criteria: A Comparison of Three Cut Score Methods. *Psychological Assessment*. *23*(2), 354-363.doi: 10.1037/0021745

Dowdy, E., Twyford, J., Chin, J., DiStefano, C., Kamphaus, R., Mays,K. (2011). Factor structure of the BASC-2 Behaviroal and Emotional Screening System Student Form. Psychological Assessment 23(2) 387.doi:10.1037/a0024000

Dowdy, E., Dever, B., DiStefano, C., & Chin, J. (2011). Screening for emotional and behavioral risk among students with Limited English Proficiency. *School Psychology Quarterly, 26*(1), 14-26. doi:10.01097/a0022072

Dowdy, E., Ritchey, K., & Kamphaus, R. (2010) School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health, 2*(4), 166-176. Doi:10.1007/s12310-010-9036-3

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

Epkins, C., & Meyers, A. (1994). Assessment of childhood depression, anxiety, and aggression: Convergent and discriminant validity of self-, parent-, teacher-, and peer-report measures. *Journal of Personality Assessment, 62*(2), 364-381. doi:10.1207/s15327752jpa6202_16.

Glover, T., & Albers, C., (2007). Universal Screening for Enhanced Educational and Mental Health Outcomes. *Journal of School Psychology, 45*(2), 117-135. doi:10.1016/j.jsp.2006.05.005.

Greenberg, M.T., Domitrovich, C., & Bumbarger, B. (2000). Preventing mental disorder inschool-aged children: A review of the effectiveness of prevention programs. Report submitted to The Center for Mental Health Services (SAMHSA), Prevention Research Center, Pennsylvania State University (obtainable at http://www.psu.edu/dept/prevention/)

Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J., Richardson, R., et al. (2006). The equivalence of the Child Behavior Checklist/1 1/2-5 across parent race/ethnicity, income level, and language. *Psychological Assessment, 18*(3), 313-323. doi:10.1037/1040-3590.18.3.313.

Guzman, M., Jelinek, M., George, M., Hartley, M., Squicciarni, A., Canenguez, K., & Murphy, J. (2011). Mental health matters in elementary school: First-grade screening predicts fourth grade achievement test scores. *European Child & Adolescent Psychiatry, 20*(8), 401-411. doi:10.1007/s00787-011-0191-3

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle(Ed.), Structural equation modeling: Concepts, issues, and applications. Thousand Oaks, CA: Sage Publications.

Hui, H. C., & Triandis, H. C. (1985).Measurement in cross-cultural psychology.*Journal of Cross-Cultural Psychology, 16,* 131–152.doi:10.1177/0022002185016002001.

Husky, M., Kaplan, A., McGuire, L., Flynn, L., Chrostowski, C., & Olfson, M., (2011). Identifying Adolescents at Risk through Voluntary School-Based Mental Health Screening. *Journal of Adolescence, 34*(3), 505-511.

Jastrowski Mano, K., Hobart Davies, W., Klein-Tasman, B., & Adesso, V., (2009) Measurement Equivalence of the Child Behavior Checklist among Parents of African American Adolescents. *Journal of Child & Family Studies, 18*(5), 606-620.doi:10.1007/s10826-009-9263-0

Kamphaus, R.W. (2011)*Screening, Assessment, and Intervention – A Comprehensive Approach to Prevention and Treatment of Emotional and Behavior Disorders*. Washington, District of Columbia, US: American Psychological Association (APA).

Kamphaus, R. W., & Reynolds, C. R. (2008). *BASC-2Behavioral and Emotional Screening System.* Minneapolis, MN: Pearson.

Kataoka, S. H., Zhang, L., & Wells, K. B. (2002). Unmet need for mental health care among U.S. children: Variation by ethnicity and insurance status. *American Journal of Psychiatry, 159,* 1548–1555. doi:10.1176/

Knight, G., Little, M., Losoya, S., & Mulvey, E. (2004). The Self-Report of Offending among Serious Juvenile Offenders: Cross-Gender, Cross-Ethnic/Race Measurement Equivalence. *Youth Violence and Juvenile Justice*, *2*(3), 273-295. doi:10.1177/1541204004265878.

Kuperminic, G.P., Leadbeater, B. J., & Blatt, S. J. (2001). School social climate and individual differences in vulnerability to psychopathology among middle school students.*Journal of School Psychology, 39*(2), 141-159. doi:10.1016/50022-4405(01)00059-0

Lau, A., McCabe, K., Yeh, M., Garland, A., Hough, R., & Landsverk, J. (2003). Race/ethnicity and rates of self-reported maltreatment among high-risk youth in public sectors of care.*Child Maltreatment*, *8*(3), 183-194. doi:10.1177/1077559503254141.

Martin, A. J., & Debus, R.L. (1998). Self-reports of mathematics self-concept and educational outcomes: The roles of ego-dimensions and self-consciousness. *British Journal of Educational Psychology, 68*(4), 517-535.

Mays, K. (2008). Examining the Predictive Validity of Screening Practices. Washington, District of Columbia, US: American Psychological Association (APA).

Merrell, K. W., McClun, L.A., Kempf, K.G., & Lund, J. (2002). Using self-report assessment to identify children with internalizing problems: Validity of the Internalizing Symptoms Scale for Children. *Journal of Psychoeducational Assessment, 20*(3), 223-239. doi:10.1177/073428290202000301

Moitabi, R., & Olfson, M. (2008). Parental Detection of Youth's Self-Harm Behavior. *Suicide & Life-Threatening Behavior, 38(*1), 50-73.

Moretti, M., Fine, S., Haley, M., Marriage, K. (1985). Childhood and Adolescent Depression: Child-report versus Parent-report Information. *Journal of the American Academy of Child Psychiatry, 24*(3), 298-302.doi:10.1016/S0002-7138(09)61090-6.

Muthén, L. ,& Muthén, B. (2004). *Mplus user's guide* (3rd ed.).Los Angeles: Muthén & Muthén

Ollendick, T., Oswald, D., & Francis, G. (1989).Validity of teacher nominations in identifying aggressive, withdrawn, and popular children. *Journal of Clinical Child Psychology, 18*(3), 221-229. doi:10.1207/s15374424jccp1803_4.

Sangwon, K., Senck-Ho, K., & Kamphaus, R. (2010). Is Aggression the Same for Boys and Girls? Assessing Measurement Invariance With Confirmatory Factor Analysis and Item Response Theory. *School Psychology Quarterly, 25*(1), 45-61. doi:10.1037/a0019768

Sharp, C., Croudace, T. ,Goodyer, I. M. & Amtmann, D. (2005) The Strength and Difficulties Questionnaire: Predictive validity of parent and teacher ratings for help-seeking behaviour over one year. *Educational and Child Psychology*. *22*(3), 28-44.

Stapleton, L., Sander, J., & Stark, K. (2007). Psychometric Properties of the Beck Depression Inventory for Youth in Sample of Girls. *Psychological Assessment*. *19*(2), 230-235.doi:10.1037/1040-3590.19.2.230

Tepper, P, Liu, C., Guo, C., Zhai, J., Liu, T., Li, C. (2008). Depressive symptoms in Chinese children and adolescents: Parent, teacher, and self-reports. *Journal of Affective Disorders*, 111(2/3), 291-298.doi:10.1016/j.jad.2008.03.013

Waters, E., Stewart-Brown, S., & Fitzpatrick, R. (2003). Agreement between adolescent self-report and parent reports of health and well-being: results of an epidemiological study. *Child: Care, Health & Development, 29*(6), 501-509. Retrieved from CINAHL Plus with Full Text database.