LEVEL OF AGREEMENT, PREDICTIVE ABILITY, AND RATINGS OF POSITIVE

BEHAVIOR OF MULTIPLE INFORMANTS ON THE BASC-2

by

UMA PARAMESWARAN DORN

(Under the Direction of Linda Campbell, Ph.D.)

ABSTRACT

Five to nine percent of children and adolescents have a serious emotional disorder and of those only 21 percent receive services. Early and accurate diagnosis can increase positive outcomes for these children (U.S. Surgeon General's Report, 1999). Behavior checklists are standard assessment tools used in the evaluation of children and adolescents for the diagnosis of various disorders. This study focuses specifically on the Behavior Assessment for Children, Second Edition (BASC-2) in addressing factors that would assist in the interpretation of the ratings of multiple informants. In particular, the following factors were considered: the level of agreement between informants, the ability of informants' ratings in the prediction of the final diagnosis, the difference in parents' and teachers' ratings on positive versus problem behavior and the differences in informants' ratings based on the child's gender.

Data was gathered from 100 evaluations that used the BASC-2 from a university based outpatient clinic that conducted assessments on children and adolescents ages 6 to 17. Ratings for Internalizing, Externalizing, Attention Problems, Adaptive Skills and Behavior Symptom Index were explored across multiple informants (mothers, fathers,

and two teachers). Given the sample size and the percentage of missing data, data was analyzed using multiple imputation techniques. Findings suggest that level of agreement between mothers and fathers are higher than the level of agreement between two teachers across all the scales except Externalizing. On the Externalizing scale, teachers' ratings had higher levels of agreement. Furthermore, results suggest that the level of agreement between informants is lower across settings (home vs. school). Differences in agreement were found for Adaptive Skills and Internalizing with parents ratings these scales higher than teachers. Results for informants' predictive ability indicated that mothers' ratings are the best predictors attention-related diagnosis and fathers' ratings for internalizing related diagnosis. No differences were found between parent and teacher ratings of positive and problem behavior. In examining gender differences, parents and teachers rated boys higher than girls for externalizing behavior, attention problems, and total problems; whereas, parents and teachers rated girls higher than boys for positive behavior. Implications and directions for future research are discussed.

LEVEL OF AGREEMENT, PREDICTIVE ABILITY AND RATINGS OF POSITIVE

BEHAVIOR OF MULTIPLE INFORMANTS ON THE BASC-2

by

UMA PARAMESWARAN DORN

B.A., Berry College, 2000

M.S., Georgia State University, 2006

Ed.S., Georgia State University, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2011

LEVEL OF AGREEMENT, PREDICTIVE ABILITY, AND RATINGS OF POSITVE

BEHAVIOR OF MULTIPLE INFORMANTS ON THE BASC-2

by

UMA PARAMESWARAN DORN

Major Professor: Linda Campbell

Committee: V. Gayle Spears
Alan Stewart
Pamela Paisley

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2011

DEDICATION

To my children:

Leia Vani, Aurea Rani, and Aidan Mahesh

This is my prayer for you:

May today there be peace within. May you trust that you are exactly where you are meant to be. May you not forget the infinite possibilities that are born of faith in yourself and others. May you use the gifts that you have received, and pass on the love that has been given to you. May you be content with yourself just the way you are. Let this knowledge settle into your bones, and allow your soul the freedom to sing, dance, praise and love...

~ Mother Teresa

ACKNOWLEDGMENTS

First and foremost to *the One who put the light in me...I just might bend but won't break as long as I can see Your face...I'm still a dreamer, a believer; You are the answer, the redeemer...But on Your shoulder I am free to be me*

My family: To my parents (Kala & Parameswaran) for being a great source of support (in so many ways) and for their patience and kindness (beyond what was expected) during this arduous process of getting a PhD. I want to thank my sister Priya for all her advice, support, picking up the kids, late night editing…etc., could not have done this without you. To my youngest sister, Hemalatha, I am glad you are home to share this with me. To Marian for your love and support.

To my friends: Jess, Shaun, Heather, Yolanda and Alessandra for carrying me through this time – you have been the ones to hear it all and who have continued to encourage me, believe in me and waited for me (to be done!).

To Dr. Linda Campbell –Thanks for taking me in the first summer, guiding me through and pointing me in the right direction. I also want to thank you for instilling my love for assessment (yes...I love assessment).

To Dr. Gayle Spears – You have been my constant in the midst of the chaos that seemed to surround me. I have learned so much from you about being true to yourself and being balanced in life. You have been the voice of reason, comfort and peace, thank you!

To Dr. Pam Paisley and Dr. Jenny Penney Oliver – I am so glad to have had the chance to work with the two of you. I will treasure the experience, as I learned from watching the two of you that work can and should be fun!

To Dr. Alan Stewart – Thanks for teaching me to love research and appreciate stats (I am still not sure I enjoyed it). I do want to thank you for all the help you provided in getting me through the dissertation.

To Dr. Joanna White – For being the inspiration to begin this journey, for believing in me and encouraging me to keep going.

Lastly to Dom, my sojourner – you have always been there for me and have never doubted my dreams (no matter how crazy they might be). I am grateful that you have allowed this journey to strengthen our relationship in a new way.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Overview of the Proposed Study

Behavior checklists in psychological evaluations are integral in the diagnosis of socio-emotional disorders in children and adolescents. A standard of practice is to use multiple informants (mothers, fathers and teachers) across multiple contexts when rating problem behaviors. The present study will examine the level of agreement and predictive ability between multiple informants on the Behavior Assessment System for Children, Second Edition (BASC-2). In addition, the study will investigate the difference in perceptions of adaptive skills compared to problem behavior across multiple informants. Lastly, the study will explore the influence of the child's gender on informants' responses.

Background and Context

A distinctive role of psychologists is to use assessments in the diagnosis of socio-emotional disorders, focusing up to 15 percent of their professional work in assessment. (Sayette, Mayne & Norcross, 2010). Although, historically assessment has been a focus of clinical psychology, research now shows that counseling psychologists are comparable to clinical psychologists in using assessments (Sayette et al., 2010). Counseling psychology's relationship to assessment began in the area of career and vocational psychology (Bentz, 2008). The backdrop for counseling psychology has been the use of assessment tools to assist college students in their academic and career paths (Brown &

1

Rector, 2008). As counseling psychology has expanded its use of assessments, the field

continues to emphasize a strengths-based perspective and infusing multicultural and

social justice principles (Lichtenberg, Goodyear, & Genther, 2008).

The Multicultural Guidelines (2003a) from the American Psychological

Association (APA) details culturally sensitive practices in the use of assessments. The

guidelines suggest that psychologists are cautious in the selection, administration, and

interpretation of assessments based on the individual differences of the client. More

importantly, psychologists need to be aware of the limitations of the instruments and any

test bias that may influence the outcome of the assessments. Similarly, counseling

psychology's emphasis on social justice encourages psychologists to address disparities

that exist among marginalized groups both at systemic and individual levels (Speight &

Vera, 2008). More relevant to assessment, psychologists are to be aware and

acknowledge their power in assessing and diagnosing and look across multiple levels to

tackle inequities among groups (Watts, 2004).

One of these multiple levels that directly affect the use of assessments and more

specifically behavior checklists is managed care and its influence on diagnosis. The

continued drive for efficient and cost-effective services placed by managed care creates

restrictive parameters for psychologists in assessing children and adolescents (Cooper &

Gottlieb, 2000). For example, in the diagnosis of Attention Deficit Hyperactivity

Disorder (ADHD), if it is found that children have symptoms of ADHD based on initial

screening using behavior checklists, some managed care companies will not reimburse

for full evaluations (which can include intelligence and achievement testing, behavior

checklists, interviews with parent and child and behavioral observations of child). Not being able to complete full evaluations may prove to be a disservice to children and adolescents, given that ADHD symptoms can be found across other disorders such as anxiety and depression (Schulz, 2002). Psychologists, therefore have to make efforts to ensure that children and adolescents are evaluated comprehensively and accurately when diagnosing socio-emotional disorders.

Significance of the Problem

APA's Task Force on Psychology's Agenda for Child and Adolescent Mental Health (TFPACAMH) stated in the executive summary of the report (2003) that the state of child and adolescent mental health "constitutes a public health crisis for our nation" (p.3). Prevalence rates of children and adolescents with mental health concerns vary from source to source. According to the U.S. Surgeon General's Report (1999), one in five children has a mental health disorder. The National Institute of Mental Health reported that 13 percent of children ages 8 to 15 had at least one mental disorder (National Health and Nutritional Examination Survey, 2004). As reported in the National Health Interview Survey (2007), slightly above 5 percent of children were reported by their parents as having serious emotional difficulties, attentional difficulties or problems getting along with others. In other prevalence research, depressive disorders were found in 5.2 percent of children and adolescents, anxiety disorders were found in 8 percent of youth and ADHD was found in 4.5 percent of youth. In addition to these emotional difficulties, these youth are less likely to be diagnosed at the onset of the disorder delaying effective treatment (National Health Interview Survey, 2007).

Children and adolescents with social, emotional and behavioral problems are at higher risk of developing more serious disorders later in life (Najman et al., 2008). For these youth if these disorders are left untreated the risk of involvement with the juvenile justice system increases (Kessler, Avenevoli, & Merikangas, 2001). Studies show 66 percent of boys and almost 75 percent of girls in the juvenile justice system with at least one mental disorder (President's New Freedom Commission on Mental Health Final Report, 2003). The effects of these problems are far reaching, impacting families and schools at systemic levels (O'Connell, Boat & Warner, 2009). Furthermore, the burden of care for youth and their families across communities makes this a vital issue to be addressed.

Families impacted by children and adolescents with socio-emotional disorders tend to experience higher levels of stress (Dababnah & Cooper, 2006). These families are also encumbered with identifying services and resources to effectively treat their children. Families who experience difficulties in dealing with the problems of their children are more likely to need support from the mental health system. Therefore, family-based services continue to be the cornerstone of mental health treatment of youth (Dababnah & Cooper, 2006). Adding to the complexities of the mental health system, parents also have to navigate the educational system, intensifying their stress.

The educational system equally shares the yoke of identifying and treating children to provide access to education for youth with emotional and behavioral disorders. More often the schools are the initial referral source for the child's entry to mental health services. A national report on special education showed that 6.7 million

students received special education services, accounting for 14% of the student

population (Center for Mental Health in Schools at UCLA, 2008). In educational settings,

teachers are central to the identification of problem behaviors especially when it becomes

difficult to manage in the classrooms.

Given the critical need for early identification of socio-emotional disorders in

youth, schools and mental health professionals require assessment tools that are effective

in diagnosing these disorders. Most professionals use behavior checklists in their

assessment of socio-emotional disorders such as ADHD, Anxiety, and Depressive

disorders. These checklists are sometimes used as stand-alone screeners or as part of

comprehensive evaluations. One of the critical benefits of behavior checklists is that they

provide a brief overview of symptoms that the child may be experiencing in various

contexts (i.e. home, school, etc.). Many of the checklists also match Diagnostic Statistical

Manual-IV-TR (2000) diagnostic criteria, which makes it easier for professionals to

diagnose. Both a benefit and complicating factor of using these checklists is the ability to

have multiple informants such as parents and teachers who may not always provide

consistent information across contexts.

<div align="center">**Purpose of the Study**</div>

Level of Agreement

In assessing socio-emotional disorders in children and adolescents, evaluators

have continued to use behavior reports from multiple informants (Merrell, 2008). Parents

and teachers, having the most access to children and adolescents, are able to report

behaviors that are interfering with academic and social functioning. Research shows that

there is low agreement between multiple informants across multiple settings (i.e. home vs. school) and moderate levels of agreement between informants in similar settings (i.e. mother and father in the home) (Duhig, Renk, Epstein, & Phares, 2000). Although there is low to moderate levels of agreement, professionals continue to use multiple informants. In using multiple informants across checklists evaluators agree 1) that each informant adds to the clinical picture of the child; (2) that each informant has a different perception that is reflected in their reports; and 3) finally a child may behave differently in different contexts (Grietens et al. 2004). Other studies have found that parents and teachers agree on externalizing behavior, but differ in reporting internalizing behavior (Woo et al., 2007; Keiley, Lofthouse, Bates, Dodge, & Pettit, 2003.). Furthermore, across research studies, the findings on level of agreement across raters have been inconsistent due to differences in types of participants, disorders being studied, the instruments and scales used and the types of data analyses (Grietens et al., 2004).

Predictive Ability

Much of the research on the predictive ability of teacher and parent ratings has been conducted on specific disorders such as ADHD. Based on a study by Tripp, Schaunghency, and Clarke (2006), teachers' ratings were stronger in specificity and accuracy of diagnosis. However, when parent ratings were combined across multiple measures the authors found that parents were just as accurate as teachers in predicting ADHD. The authors also found that parents and teacher reports were not predictive of a diagnosis when only using a single measure. The authors in this study recommend a more

comprehensive assessment that includes clinical interviews and differential diagnosis using different informants with multiple measures.

Adaptive Skills

The Adaptive Skills subscale on the BASC-2 measures competence and skills across academic and social domains addressing positive aspects of the child. Research on parents' and teachers' perceptions of adaptive skills compared to problem behavior is limited. Parents are seen as the most reliable source to report adaptive skills (Loeber, Green & Lahey, 1990; Phares, 1996). Related to adaptive skills, one study reported that teachers rated adolescents with lower positive behavior than parents (Hines & Paulson, 2006). These studies suggest that parents may be more apt to describe positive or adaptive skills than other informants.

Gender

In relation to gender, parents reported more males (6%) than females (4%) ages 4-17 as having serious emotional or behavioral difficulties (National Health Survey 2007). In other studies pertaining to the child's gender and informant ratings, conflicting results were found. For example, studies found that gender did not affect the way informants rated the child (Achenbach, McConaughy, & Howell, 1987; Konold, Walthall, & Pianta, 2004). However, in Duhig et al. (2000) the authors found that collectively gender influenced the agreement between mothers and fathers, but was not an influence when there were discrepancies between mothers and fathers.

Based on the research, it is accepted that there are higher levels of agreement in the same settings (home/school); and lower levels of agreement across settings (home vs.

school) (Verhulst & Akkerhuis, 1989; Youngstrom, Loeber, & Stouthamer-Loeber, 2000; Merrell, 2008). The ability to predict various disorders varies based on whether the disorder includes behaviors that are considered internalizing or externalizing. Research shows that informants are more likely to agree and identify externalizing behaviors than internalizing behaviors (Merrell, 2008). In examining adaptive skills, the little research that exists gives evidence to parents rating their children with more adaptive skills than any other informants (Hines & Paulson, 2006). Lastly, the research on gender interaction on the informants' responses had mixed results with some studies finding no gender effects and other studies that have found some differences specifically related to informants' rating boys with more externalizing behaviors compared to girls (Duhig et.al, 2000).

It is important to note a majority of this research has been conducted on the Achenbach System of Empirically Based Assessment (ASEBA) as evidenced by over 7000 publications on the ASEBA bibliography database (ASEBAbib.com) compared to approximately 500 publications for the BASC and BASC-2 (PsychINFO). ASEBA instruments are commonly referred to as the Child Behavior Checklist (CBCL) and widely used for identifying emotional and behavior disorders in children (Merrell, 2008; Hughes & Melson, 2008). The BASC-2 and CBCL are both assessment tools that have similar constructs and use self-report, parent and teacher reports. Given that the extensive research has been done on the CBCL and limited research on the BASC-2, the examination of this study will focus on the BASC-2.

This study proposes to examine Parent and Teacher Reports of the BASC-2 in the diagnosis of socio-emotional disorders in children and adolescents. More specifically, the present study aims to examine the agreement between both parents and teachers regarding their reports of behaviors on the BASC-2 for the child. Also, this study will evaluate the ability of each informant to predict a diagnosis of a socio-emotional disorder. Ratings of positive behavior and problem behavior will be compared across multiple informants. In particular, the study will examine the following subscales on the BASC-2: Adaptive Skills, Internalizing, Externalizing, Attention Problems and Behavior Symptoms Index. In studying 1) the level of agreement between multiple informants 2) the ability of each informant to predict the final diagnosis 3) the difference in informants' ratings of positive behavior compared to problem behavior and 4) the gender interaction in the informants' ratings, the researcher hopes to obtain information that will aid in developing effective identification tools in diagnosing socio-emotional disorders.

Research Questions

Since research is limited on the BASC-2 and given the inconsistent findings across other studies of behavior rating scales, this study will focus on the BASC-2 to examine ratings of multiple informants (mother, father, and teachers) across Internalizing, Externalizing, Adaptive Skills, Behavioral Symptom Index and Attention Problems subscales in 6-to 17-year old children and adolescents from a clinical sample. The study will address the following research questions:

1. What is the level of agreement between ratings of informants (mothers, fathers, and teachers) across subscales of the BASC-2?

9

2. How do different informants differentially and/or similarly predict the final diagnosis (including having no diagnosis)?

3. Do informants' perceptions differ on positive behavior compared to problem behavior?

4. Does gender of the child interact with informants' ratings?

General Hypothesis

Based on current research using behavior checklists the researcher predicts that the level of agreement between the parents will be moderate to high given parental perceptions of their children and perceptions across internalizing and externalizing domains. Parents may report behaviors that are more likely than other informants to predict the final diagnosis of a socio-emotional disorder. Moreover, teacher reports will have higher levels of agreement with parents on externalizing behaviors. Teacher reports also will be more predictive of a final diagnosis related to externalizing behaviors than parent reports. There will be differences between how informants' rate positive behavior; overall, parents will report more positive behavior than teachers. In regards to gender, it is hypothesized that informants will report more problem behaviors in boys than in girls.

Definitions and Key Terms

Socio-emotional disorders: For the purposes of this study social, behavior and emotional problems as met by Diagnostic and Statistical Manual of Mental Disorders, 4th. Edition, Text Revision (DSM-IV-TR) (2000) criteria for Attention-deficit and Disruptive Behavior disorders, Mood Disorders, Anxiety Disorders and Adjustment Disorders will be used.

Behavior checklists/rating scales: standardized assessment tools used to gather raters

perception of behavior.

Informants: Persons who have completed rating scales on children and adolescents

including mothers, fathers and teachers.

Level of agreement: The inter-correlation` between the informants' scores on the

behavior checklists. For this study it is the agreement between informants (mothers,

fathers and teachers) on the various subscale of the BASC-2: Internalizing, Externalizing,

Adaptive skills, Behavioral Symptom Index and Attention Problems.

Internalizing: Behaviors that are reflective of internal disturbances such as being

withdrawn, depressed, lonely, anxious, etc. (Smith, 2007).

Externalizing: Behaviors that are reflective of "acting out" such as being aggressive,

impulsive, defiant etc. (Smith, 2007).

Attentional problems: Behaviors that include inability to sustain attention, difficulty

concentrating, easily distracted, fidgety, has difficulty remaining in seat etc. (DSM-IV-

TR, 2000).

Adaptive skills: Focuses on positive behavior and competencies in skills across

communication, daily living skills, academic and social domains (Reynolds &

Kamphaus, 2004). This term differs from "adaptive behavior" which is used in research

with the developmentally delayed and mentally retarded population (Szatmari, Archer,

Fisman, & Streiner, 1994).

Narrow band scales: Scales that measure specific disorders such as ADHD, depression

and anxiety (Merrell, 2008).

Broad band scales: General purpose scales that measure wide range of behaviors such as the CBCL and BASC-2 (Merrell, 2008).

ASEBA: Achenbach System of Empirically Based Assessment commonly referred in professional practice as the CBCL is a broadband measure of problem behaviors and competencies and includes Parent, Teacher and Self-Reports. The CBCL is the parent version of the instrument and the TRF is the teacher version (Merrell, 2008).

BASC-2: Behavior Assessment System for Children is a multidimensional assessment tool designed to assess adaptive and problem behaviors. For the purposes of this study the Teacher and Parent Reports will be used (Reynolds & Kamphaus, 2004). The first edition of the BASC was developed in 1992 and the BASC-2 was developed in 2004.

CHAPTER 2

LITERATURE REVIEW

Uses of Behavior Checklists

Recognizing problem behaviors across various contexts is crucial in the early

identification of socio-emotional disorders. Behavior checklists have been used across

home and schools to identify and diagnose mental disorders in children and adolescents.

There are multiple methods in assessing youth including direct observations and clinical

interviews; however, the advantages of behavior checklists have made it an increasingly

popular assessment method in comprehensive evaluations and as screeners (Hughes &

Melson, 2008). The advantages of using behavior rating scales include the cost

efficiency; increased reliability and validity compared to other methods; and the use of

multiple informants across home and school. There are also some limitations to using

behavior checklists including error variances across instruments, among informants, and

between home and school. Other disadvantages include respondent bias where the

informant may minimize or exaggerate symptoms (Merrell, 2008). Behavior rating

scales are often utilized in: a) screening and early identification of mental disorders; b)

gathering "aggregate" data from multiple informants to get the best picture of the child;

and c) to evaluate or monitor progress during and after treatment (Merrell, 2008).

Behavior checklists are used to identify various general purpose problem

behaviors (broad band scales) or designed for a specific disorder (narrow band scales)

such as Attention Deficit Hyperactivity Disorder (ADHD) (Merrell, Streeter, Boelter,

2001). The BASC-2 is used to identify clusters of symptoms across multiple disorders. This evaluation tool assesses symptoms of common disorders among children and adolescents, including Anxiety, Depression, Oppositional Defiant Disorders, Conduct Disorders, Autism, and ADHD.

Socio-emotional Disorders

Parents reported several conditions related to emotional and behavioral disorders, according to a National Health Interview Survey (2008). The survey found that 53.5 percent of children with emotional and behavioral problems had at one time been diagnosed with an attention disorder. Additionally, 43.5 percent of children had depression or anxiety problems and 55.7 percent had a behavioral or conduct problem. These problems impact children and adolescents in their homes, schools and communities. Socio-emotional problems that are left untreated can contribute to severe negative consequences in the future. These consequences range from suicide, involvement with the justice system, and substance abuse (APA, 2003b).

The socio-emotional disorders that are the focus of this study are common disorders among children and adolescents. Broader categories of *Internalizing, Externalizing* and *Attention Disorders* will be used to classify the socio-emotional disorders. The following is a brief description of each broad category of disorders and current research related to the use of behavioral checklists in assessing these disorders.

Internalizing Disorders

Internalizing disorders include problems related to disturbances that are internal in nature. The DSM-IV-TR (2000) categories for Mood Disorder and Anxiety Disorders can be classified as internalizing disorders.

Mood Disorders

Mood disorders are divided into Depressive Disorders and Bipolar Disorders which will be the focus of the current study. Depressive disorders primarily found in adolescence are found in 2.5 percent of children and 8.3 percent of adolescents (Birmaher et al. 1996). In a NIMH study prevalence rates of 6 percent were found for depression in 9 to 17-year-olds within a 6-month period with up to 4.9 percent meeting criteria for major depression. Studies have also found that mood disorders tend to persist and recur in adulthood (Kessler, Avenevoli, & Merikangas, 2001). In addition, depression coincides with other mental disorders such as Anxiety, Disruptive Behavior Disorder and Substance Abuse (Angold & Costello, 1993).

Narrow band scales are typically used to measure depression such as the Children's Depression Inventory (CDI) or the Beck's Depression Inventory for older adolescents. The BASC self-report has been shown to measure similar constructs as the CDI (Salvia, Ysseldyke, & Bolt, 2009). Depression is usually measured by self-reports as it is difficult for parents and teachers to rate internalizing behavior. Depression also manifests itself differently in children with more "acting out" behaviors that may not include withdrawn behavior. Much of the research has been conducted on self-reports, although in one study it was reported that there was low congruence between teachers, parents and self-reports measuring for depression on the BASC (Brook, 1996).

15

Additionally, this study found that with more severe depression parent and teacher reports did not agree with youth self-reports.

Dysthymia is characterized by chronic depression that is milder in course compared to Major Depressive Disorder (MDD). Symptoms are similar in both disorders with feelings of hopelessness, low energy, changes in appetite and changes in sleep. However, in Dysthymia there is a low grade depression that persists whereas MDD can fluctuate in intensity (Carr, 2008). Prevalence rates of Dysthymia have been found from 1.6 to 8% in adolescents. Research on the CBCL has found that both the Parent and Youth reports are effective in screening for Major Depression and Dysthymia (Ferdinand, 2008). The BASC-2 has also been found useful in screening for depression (Carr, 2008).

The diagnosis of Bipolar Disorder (BPD) has increased 260% from 1994-2001 (Younstrom, Findling, Youngstrom, & Calabrese, 2005) and is diagnosed in 3.9% of the pediatric population. It is difficult to assess for BPD as symptoms overlap in other disorders such as depression and ADHD. Although the BASC-2 assesses for manic symptoms, the research on the using behavior checklists for BPD has been primarily limited to the CBCL. The CBCL has a specific scale for assessing Juvenile BPD and has been found effective in screening for such using Parent Reports (Farone, Althoff, Hudziak, Monuteaux, & Biederman, 2005).

Anxiety Disorders

Some of the more common anxiety disorders in children and adolescents are Generalized Anxiety Disorder, Separation Anxiety Disorder, Phobias, Panic Disorder, Obsessive-Compulsive Disorder and Post-traumatic Stress Disorder. These disorders are

characterized by intense fear, worry and nervousness (DSM-IV-TR, 2000). Anxiety

occurs in 13 of every 100 children from ages 9 to 17. Females are more likely to develop

an anxiety disorder than males. Anxiety disorders also co-occur with either another

anxiety disorder or other mental disorders (National Health Interview Survey, 2007).

Similar to depression many of the instruments for anxiety use narrow band self-

report scales such as the Children's Manifest Anxiety Scale (CMAS) and the Beck's

Anxiety Scale for older adolescents. However, the BASC-2 has been shown to be

effective in identifying anxiety in children and adolescents when compared to CMAS

(Salvia, Ysseldyke, & Bolt, 2009). In comparing self-reports with parent and teacher

reports on social skills and anxiety there was low agreement between the informants on

the BASC-2. The authors point out that both adult raters and child reports appeared to be

valid, however the difficulty is deciding which to use and how to incorporate the

discrepant information (Salvia et al., 2009). In a similar study comparing parents and

child reports, parents reported higher levels of anxiety than their children. There were

low levels of agreement between parent and self-reports (Martens, 2006).

Generalized Anxiety Disorder (GAD), commonly found in children and

adolescents is characterized by excessive and uncontrollable worry that persists for at

least 6 months (Silverman & Ollendick, 2005). Studies on the CBCL have found that the

Internalizing scale, specifically the Anxious/Depressed subscale is predictive of GAD in

children and adolescents (Petty et al., 2008; Aschenbrand, Angelosante, & Kendall,

2005). BASC-2 has been found effective in screening for anxiety in children with Autism

(Volker et al., 2010; Kimel, 2009).

17

Externalizing Disorders

Externalizing disorders include problems related to disturbances that are considered "acting out" behaviors. The DSM-IV-TR (2000) categories for Attention Deficit and Disruptive Behaviors can be classified as externalizing disorders. However, for this study a further delineation will be made between Disruptive Behaviors (as externalizing disorders) and ADHD (attention disorders).

Disruptive Behavior Disorders

Disruptive Behavior Disorders are characterized by negative, hostile or defiant behaviors. Disruptive Behavior Disorder includes Conduct Disorder and Oppositional Defiant Disorder (ODD). Conduct disorders, one of the most frequently diagnosed disorders in mental health settings, range from 1 percent to more than 10 percent. Higher percentage of males than females are diagnosed with Conduct disorder. ODD in youth has been found between 2 percent and 16 percent depending on the source of information (National Health Interview Survey, 2008). Broad band scales such as the BASC-2 are often used in identifying youth with Conduct and Oppositional Defiant Disorders with success in distinguishing between cases and non-cases (Muratori, Salvadori, Picchi, & Milone, 2004; Scholte, Van Berckelaer-Onnes & Van der Ploeg, 2008).

Attention Disorders

Attention Deficit characterized by attention related problems such as inattention, hyperactivity and distractibility. Specific disorders relevant to this study are Attention Deficit Hyperactivity Disorder and its subtypes (Hyperactive, Inattentive and Combined). ADHD has been found in 3 to 7 percent of school-age children (DSM-IV-TR, 2000) and

8 percent in children 3-17 (National Health Interview Survey, 2008) with a higher percentage found in males (11 percent for males compared to 4.8 percent of females).

Many behavior checklists have an emphasis in the assessment of ADHD (Merrell, 2008). In a comparison study conducted by Vaughn, Riccio, Hynd & Hall (1997), the authors found that between the BASC and the Child Behavior Checklist (CBCL), both instruments were equally effective in identifying ADHD from those who do not meet criteria for ADHD. It was discovered that the BASC Teacher Report was more effective in identifying youth who did not meet criteria. An additional finding was that both the Parent and Teacher forms of the BASC were able to better identify subtypes of ADHD than the CBCL.

Collett, Ohan, & Myers (2003) conducted a 10-year review of ADHD rating scales and put forth some recommendations for mental health professionals to consider. They suggested when using narrow band scales that are based on the DSM-IV-TR (2000) such as the Attention Problems scale on the BASC-2, it is important to remember that high scores on these scales may not necessarily mean that there is evidence of a diagnosis. In general ratings scales are scores based on the perceptions of the responders and not the existence of the behavior. The face validity on these scales makes it easy for respondents to rate based on what they believe (the child has ADHD or does not). Another finding is that most ADHD scales measure the number and severity of symptoms, but not the presentation of symptom which may vary across multicultural factors (age, gender, race, SES, etc.).

Given that more often boys are diagnosed with ADHD than girls, the authors in this study point out their concerns. First the DSM-IV-TR (2000) uses age as a criteria, but not gender, which makes scales that use a combined normative group for boys and girls a potential problem. Cut-off scores for these norm groups will result in more boys meeting ADHD criteria than girls. They suggest that gender-specific norm groups be used for males and females in its normative data which the BASC-2 includes. Lastly, it is important to note that a majority of the scales were developed based on data from males which may not accurately represent symptoms in females.

The research on ADHD and comorbidity, has found that children with ADHD often have other coexisting disorders that affect diagnosis and treatment (Smith, Barkley & Shapiro, 2007). ADHD is commonly found with Depression, Bipolar Disorder, Anxiety, Tic Disorders and Oppositional Defiant Disorder and Conduct Disorders (Spencer, 2006; Fischer, Barkley, Smallish & Fletcher, 2004). These disorders often have similar symptoms of ADHD making it difficult to diagnose (Barkley, 2006). Furthermore, comorbidity is associated with increased negative outcomes educationally and socially (Larson, Russ, Kahn, and Halflon, 2011). According to Larson et.al., (2011), they found that 46% of children with ADHD had a Learning Disability compared to 5% of non-ADHD children. Broadband behavior checklists like the CBCL and the BASC-2 are designed to assess for these disorders to assist in differential diagnosis (Merrell, 2008). Since most of the research has been based on CBCL, an overview of the instrument and the synopsis of the findings related to the variables in this study will be presented below.

Child Behavior Checklist (CBCL)

Overview of the Instrument

 Much of the research in using behavior checklists has focused on the CBCL, a broadband instrument that is similar to the BASC-2. The CBCL includes parent, teacher and self-report versions to assess competencies, problem and adaptive behaviors. The original instrument was developed in 1966 to empirically identify disorders from the DSM-I (1952). The instruments have been through many revisions with the latest in 2001 that includes DSM-oriented scales based on DSM-IV-TR (2000) and include disorders such as Oppositional Defiant Disorder. Multicultural considerations and applications in 2007 were added in to address cultural differences in the use of these instruments (Achenbach, 2010). CBCL continues to be used extensively in mental health services, schools, hospitals, HMOs, public health agencies and in training and research (Achenbach, 2010). It includes checklists across the life span from ages1 ½ to 90 years old. The CBCL includes eight syndrome scales that have T-scores cut off at 65 for the Borderline clinical range and T-score of 70 for Clinical range. The Internalizing subscale includes Anxious/Depressed; Withdrawn/Depressed; and Somatic Complaints and the Externalizing subscale includes Rule-Breaking Behavior and Aggressive Behavior (Achenbach & Rescorla, 2001). The CBCL has shown to be highly correlated with the Parent and Teacher Reports of the BASC-2 (Achenbach et.al, 2008). Research across 20 different cultures has shown that there is external validity to the CBCL (Ivanova, Achenbach, Rescorla, Dumenci, Almqvist, & Bathiche, 2007; Ginzburg, 2009).

Continued research is needed on persons of color and immigrant populations (Ivanova et al., 2007).

Informant Agreement

The CBCL has been the focus of many studies investigating informants' ratings and predictive ability. Ratings across informants vary based on the types of behavior (internalizing vs. externalizing), across types of informants (parents vs. teachers) and settings (home vs. school) (Achenbach, McConaughy,& Howell, 1987; Duhig, et al., 2000). Overall, there is low agreement between informants across settings such as home and school and moderate levels of agreement among informants in same setting such as mothers and fathers in the same home (Merrell, 2008; Achenbach, McConaughy, Howell, 1987; Christensen et.al, 1992; Duhig et.al, 2000). Multiple studies have found that there are higher levels of agreement among raters when reporting externalizing behavior (Christensen et al., 1992; Duhig, et al., 2000; Grietens et al., 2004). These same studies have also found that parents and teachers have low agreement in reporting internalizing behaviors. In regards to positive behavior versus problem behavior the limited research shows that parents are more likely to report positive behavior or adaptive skills (Achenbach & Howell, 1993; Achenbach, Dumenci, Rescorla, 2002). Studies also report that parents particularly mothers, more often report the most problem behaviors (Sawyer, Baghurst & Mathias, 1992; Lee, Elliot & Barbour, 1994).

Two meta-analytical studies on the CBCL (Achenbach, McConaughy,& Howell, 1987; Duhig, et al., 2000) were conducted examining informants ratings. The first study by Achebach et. al (1987), suggest that there are moderate levels of agreement between

mothers and fathers when assessing problem behaviors. The authors reviewed 119 studies and performed data analysis on 269 samples including reports from multiple informants. The study found that mothers and fathers correlated on average at .60. Similarly, teachers at school also correlated on average at .60. However, when looking at parents and teachers across home and school, the correlation was much lower (.28). To further examine differences in interparental agreement, Duhig et.al (2000) focused their meta-analytic study to examine the discrepancies in ratings of mothers and fathers. The authors reviewed 60 studies and 120 effect sizes. Similar to Achenbach et.al (1987), the authors found that there were moderate levels of agreement between parents. Across home and school, other studies also found low levels of agreement between teachers and parents (Kolko & Kazdon, 1993; Stanger & Lewis, 1993; Youngstrom, Findling, & Calabrese, 2003). The results of these studies give evidence to low levels of agreement between parents and teachers across home and school and moderate levels of agreement in the same setting (home/school).

Differences on levels of agreement between raters were also found across internalizing and externalizing behaviors (Duhig et al., 2000). In this meta-analysis, the authors found high levels of agreement between mothers and fathers on externalizing behavior. Recent studies found similar results where mothers and fathers had higher levels of agreement on externalizing behavior compared to internalizing behavior (Gritens et al., 2003; Salbach-Andrae, Lenz, & Lehmkuhl, 2009). Teachers and parents also had higher levels of agreement on externalizing behaviors and overall problem behaviors (Grietens et al., 2004; Woo et al., 2007; Cai, Kaiser & Hancock, 2004). Also

multiple Teacher Reports differed in their ratings of internalizing and externalizing

behaviors (Achenach et. al, 1987). The research indicates that externalizing behavior is

easily observed by raters which support higher levels of agreement between informants.

Internalizing behavior however an internal process maybe more difficult to observe and

rate these problem behaviors.

Although much of the research suggests that there are low levels of agreement

between informants from home and school, some studies have found higher levels of

agreement between informants with specific problem behaviors (Merrell, 2008).

Research on ADHD has found that there are higher levels of agreement on hyperactive

behaviors compared to inattentive behaviors (Loughran, 2003; Papageorgioui, Kalyva,

Dafoulis, & Vostanis, 2008). Consistent to the meta-analytical research, these studies still

had low levels of agreement across between teachers and parents on total problem

behaviors.

Studies have also examined the factors (i.e family environment and maternal

depression) that influence these discrepancies between informants (Christensen et al.,

1992; McFarland & Sanders, 2003). Many of the studies have focused on the effects of

maternal depression which has shown that with increased depressive symptoms, mothers

tend to over report problem behaviors in children and adolescents (Querido, Eyberg, &

Boggs, 2001, Nielsen, Vika, & Dahl, 2003; McFarland & Sanders, 2003). In a study that

examined family environment (Christensen et al., 1992), reports of mothers and fathers

differed greatly when the child came from a distressed family environment.

Predictive Ability

Research on the CBCL has shown that raters are able to accurately predict certain diagnosis such as ADHD, ODD and Conduct Disorder (Scholte, Van Berckelaer-Onnes & Van der Ploeg, 2008; Granero, Lourdes, Ezpeleta, Osa & Doménech, 2009). Scales on the CBCL specifically designed to measure these disorders are most accurate in predicting diagnoses (Tripp, Schaughency, & Clarke, 2006). Interestingly, in a longitudinal study across two age spans (8-years-old and then 16-years-old), in which parents and teachers completed the CBCL the findings suggested that teachers who reported conduct problems at age 8, were predictive of more severe externalizing problems at age 16. On the hand parents who reported emotional problems at age 8, were predictive of more severe internalizing problems at age 16 (Sourander & Helstelä, 2005). Another study found that although mothers and fathers had lower levels of agreement across problem behaviors, they were more predictive of later outcomes in children and adolescents (Hay et al.,1999; Verhulst, Dekker & van der Ende, 1997).

Gender Interaction

The two meta-analytic studies discussed earlier found contradictory results in regards to gender effects on informants' ratings. Achebach et al. (1987) found that gender of the child or parent did not influence informants' reports. Recent studies still find that gender has little to no affect on informants' ratings (Carbonaro, 2009). On the other hand, Duhig et al. (2000) found that collectively gender influenced the informants' ratings. The study found that mothers and fathers rated more externalizing behaviors for boys than girls. Other studies similar to Duhig et al (2000) found that agreement for externalizing was higher for boys than for internalizing problems (Greitens et al., 2004). In regards to

the gender of the informant, mothers reported more problems behaviors and internalizing behaviors compared with fathers (Lee, Elliot & Barbour, 1994; Sawyer, Baghurst, & Mathis, 1992).

On the whole, informants experience the child differently and children behave differently at home and school which may or may not be influenced by gender. Teachers observe other children in their classroom and are able to compare behavior; whereas parents may have less access to observe typical development. This may result in over reporting and under reporting of problem behaviors across the subscales. Researchers across studies have consented that discrepancies exists due to 1) differences in informants' perception; 2) children's behavior is different between home and school; 3) specific disorders (internalizing vs. externalizing) differ in expression and observability; and 4) other factors such as family environment, maternal depression, family violence affects ratings (Kolko & Kazdin, 1993; Treutler & Epkins, 2003; Greenbaum, Dedrick, Prange, & Friedman, 1994; Grietens et al., 2004; Buehler et al.,1997; Christensen et al., 1992). Researchers and clinicians suggest using multiple methods and assessment tools to help integrate discrepant ratings from informants to help in accurate diagnosis (Merrell, 2008).

<div align="center">**Behavioral Assessment System for Children (BASC)**</div>

Overview of the Instrument

The first edition of the BASC was developed in the early 1990's by Reynolds & Kamphaus (1992) which included parent, teacher, and self-report forms. The system also includes a structured developmental history form and a student observation form. The

BASC is used to evaluate problem behaviors in the home and at school and also assess

adaptive skills. The BASC-2 was developed in 2004 and changes included "(a) new

scales and item content, (b) improved normative samples and psychometric properties,

(c) a new item response format for the self-report measure, (d) new software and

interpretive reports, and (e) an expanded development of the Spanish-language forms"

(Waggoner, 2005 p.6316). Changes also included expanding the age range from 21 years

of age to 25 years of age. The BASC has been traditionally used for screening for

problem behaviors, risk assessment, tracking outcomes, eligibility for services and

monitoring progress in treatment. In addition to these uses according to Reynolds and

Kamphaus (2004), "the BASC-2 is also designed to evaluate various aspects of behavior

and personality, including positive (adaptive) as well as negative (clinical) dimensions

(p.234)."

The BASC-2 includes two broad categories across the Self-Report, Parent and

Teacher report. These categories are labeled as clinical scales and adaptive scales. The

subscales for each category vary based on the age of the child and the informant. The

Adaptive scales for age 6 and up include Social Skills, Study Skills (only on the Teacher

Report), Leadership, Activities of Daily Life (only on the Parent Report), and Functional

Communication. Clinical subscales are comprised of Internalizing, Externalizing and

Behavioral Symptoms. The BASC-2 includes Aggression, Conduct Problems and

Hyperactivity scales under the Externalizing subscale. The Internalizing subscale of the

BASC-2 is comprised of Anxiety, Depression, and Somatization scales. The Behavioral

Symptoms Index includes scores from Aggression, Hyperactivity, Anxiety, Depression,

Attention Problems, Atypicality and Withdrawal scales. Lastly, there is an additional

School Problems scale on the teacher report that includes Learning Problems and Study

Skills. The BASC-2 has a new scoring system that now includes (year) 7 content scales

Anger Control, Bullying, Developmental Social Disorders, Emotional Self-Control,

Executive Functioning, Negative Emotionality, and Resiliency. Descriptions of the scales

used in this study are found in Appendix A.

The cut off scores are based on converting raw scores to T-scores. On the

adaptive scales, T-scores of 30-40 are classified in the At-Risk range and below 30 for

Clinically Significant range. On Clinical scales, and Behavioral Symptom Index scale the

At-Risk range is a T-score between 60 and 70. The Clinically Significant cut off score is

70 for these scales. Internalizing and Externalizing scales have similar T-score as the

Clinical Scales for At-Risk and Clinically Significant. A summary of the BASC-2 scale

structure and scoring system is provided in Appendix B.

Strengths and Limitations of the Instrument

The BASC-2 is a complete assessment system that is effective in identifying

problem behavior in children and adolescents. It also uses a strengths-based perspective

and assesses adaptive skills. The multiple informant forms allow for a comprehensive

assessment of behavior. The instrument includes validity scales to determine respondent

bias. Additionally, the BASC-2 includes separate scales for differential diagnosis such as

anxiety and depression, unlike the CBCL which includes both disorders on the same scale

(Frick, Barry, & Kamphaus, 2009).

Some of the limitations of the BASC-2 include the number of items on the forms with the parent form comprised of 160 items. The BASC-2 may not be conducive for frequent monitoring due to the length of the instrument (Merrell, 2008). The BASC-2 also does not have a format to include qualitative details on item responses. Lastly, there is limited research on the parent form and it has not been validated for program evaluation or treatment planning (Frick et al., 2009; Salvia et.al, 2009).

Informant Agreement

Research on informant agreement has focused on CBCL. Limited research exists on the BASC-2 related to level of agreement between informants (Frick et al., 2009). In a recent study (Van Slyke, 2008) using informants the author found that teacher-teacher dyads had the highest level of agreement more specifically on externalizing domains. Parent-teacher dyads had overall the lowest correlations with higher levels of agreement on the externalizing behaviors. Regardless of ethnicity or gender, parents in this study reported more problem behaviors compared to other informants. In another study examining ratings on anxiety and depression subscales on the BASC, the author found on the anxiety subscale parents and children were more congruent than teachers when compared to child self-reports; however, on the depression the level of congruence was not affected by the type of informant (Brook, 1996). Likewise another study found that correlations between parent and teacher reports were low and correlations between youth reports were low to moderate for both teachers and parents (Neill, 2002). Similar results as the CBCL have been found on the BASC-2; with higher levels of agreement between for teacher-teacher and for mother-father especially related to externalizing behaviors and

lower levels of agreement across parents and teachers and internalizing behaviors (Van

Slyke, 2008).

Predictive Ability

The BASC-2 has been found to have strong construct validity based on DSM-IV-

TR (2000) criteria. However, to the extent that the informants' responses actually predict

diagnosis varies based on a number of factors. It has been noted in numerous studies on

the CBCL that it is more difficult for informants to identify internalizing behaviors than

externalizing behaviors making it difficult to use behavior rating scales to diagnose

internalizing disorders such as anxiety or depression (Granero et.al, 2002; Tripp,

Schaughency, & Clarke, 2006). Additionally, it has been found that agreement between

informants' ratings of severity of behaviors was higher when the behaviors were less

severe; whereas at more severe levels of anxiety and depression, the level of agreement

between informants was very low (Brook, 1996). The research on the BASC has shown

to have superior predictive ability compared to the CBCL in identifying ADHD

Combined Type and in discriminating between children with and without ADHD

(Saklofske, Andrews, Janzen, & Phye, 2002).

Gender Interaction

Research on the BASC-2 has had similar findings to the CBCL related to gender

and its influence on informants' ratings. In a recent study, the BASC-2 teacher and parent

reports were examined and findings suggested moderate levels of agreement for boys on

externalizing behavior and low level of agreement for girls (Van Slyke, 2008). Other

studies found that gender differences among informants were dependent on type of

problem behavior with more boys reported as having externalizing behavior and girls

having internalizing behavior (Gross-Tsur et al., 2006; van der Ende & Verhulst, 2005;

Sourander & Helstelä, 2005). Research on the gender of the informant has found that

mothers more often report problem behaviors than fathers (VanSlyke, 2008).

Adaptive Skills

Using a strengths-based perspective the BASC-2 includes the Adaptive Skills

scale to assess positive behavior in children and adolescents. Majority of the research has

compared informant ratings across problem behaviors, but limited research exists on

agreement among informants on adaptive skills or how informants rate adaptive skills

compared to problem behavior. In a study (Acuna, 2003) examining adaptive skills on the

BASC in emotionally disturbed children, the results yielded lower parent ratings on

Adaptive scales compared to problem behaviors. They also found that both parents and

teachers rated the emotionally disturbed children significantly lower on the Adaptability

scale than other scales. Kamphaus and his colleagues (1999) using normative data found

that the Adaptive Skills scale was an important construct in the original BASC as it

helped to identify children free from psychological symptoms. The authors grouped these

children as having protective factors from developing mental disorders. Research on

informants' ratings of adaptive skills is sparse and more research is needed in this area.

Similarities and Differences between the BASC-2 and CBCL

The BASC-2 and CBCL are both instruments that are used in the assessment of

problem behaviors and positive behaviors in children and adolescents. In the book,

Assessing Childhood Psychopathology and Developmental Disabilities (Rescorla, 2009),

a chapter is focused on comparing the BASC-2 and the ASEBA (CBCL) systems. They have similar constructs and include broad and narrow band scales (p.120). Although similar in many ways, there are significant differences between the CBCL and the BASC-2. Variation in the two systems are found in the arrangement and ratings the items, the assessment of positive functioning/adaptive skills, validation methods, the design of the problem scales, how informant bias is addressed, cut-off points for at-risk and clinical significance, research base and multicultural applications (p.121). In summarizing the main differences Rescorla offers the following conclusions regarding the BASC-2 and the CBCL (p.141). The CBCL splits up problem items and adaptive/competencies and uses a 3-level likert scale whereas the BASC-2 mixes problem and adaptive behavior and uses a 4-level likert scale. The CBCL assesses for competencies while the BASC-2 has more adaptive scale and uses descriptions of positive functioning. Based on the differences in cut-off points, the BASC-2 is more sensitive to assessing At-Risk populations and the CBCL is more sensitive to assessing Clinically Significant populations on the broadband scales. The CBCL has common scales to the teacher, parent and youth forms whereas the BASC-2 includes fewer scales that are common to the various forms. Lastly, extensive research (over 7000 publications) has been conducted on the CBCL compared to the limited research on the BASC/BASC-2 (over 500). Given these salient differences and the scarcity of literature in examining the BASC-2, the focus of this study on the BASC-2 is essential in broadening the research base.

Proposed Study

In reviewing the current literature on behavior rating scales related to the variables in this study, the following findings were consistently reported across studies. The level of agreement between informants mother-father and teacher-teacher is higher than informants across various settings (home vs. school) (Grietens et al., 2004; Achenbach, McConaughy, Howell, 1987; Christensen et.al, 1992; Duhig et.al, 2000, Brook, 1996; Neill, 2002; Van Slyke, 2008). These studies also found that the agreement between raters is higher for externalizing versus internalizing behavior. Another major finding is that predictive ability is higher for narrow band scales such as ADHD and for externalizing behavior compared to internalizing behavior and broad band scales (Saklofske et al., 2002; Scholte et al., 2008; Granero et al., 2009). In general studies show that parents report more problem behaviors than other informants (Achenbach & Howell, 1993; Achenbach et al., 2002). Given the limited research on adaptive skills, the current findings suggest that parents and teachers report more problem behaviors than positive behavior and that this positive behavior is seen as a protective factor (Kamphaus et al., 1999; Acuna, 2003). Lastly, gender interaction effects across studies have had inconsistent findings although, a number of studies have found that informants report more externalizing behavior with boys and internalizing behavior with girls (Achebach et al.,1987; Carbonaro, 2009; Greitens et al., 2004; Duhig et al, 2000).

Research on the BASC-2 has been limited even though it is widely used in clinical settings to assess emotional and behavioral disorders (Merrell, 2008). This study sought to extend the current body of research by examining the level agreement and the

predictive ability of informants on the following scales: Adaptive Skills, Internalizing, Externalizing, Attention Problems, and Behavioral Symptom Index. The current study examined how multiple informants rated Adaptive Skills compared to Behavioral Symptom Index. Lastly, the interaction between the gender of the child and informants' ratings was investigated. This study fills the void in the literature on the BASC-2 relative to the following factors: agreement in informants' ratings, predictive ability of the scales, ratings on adaptive skills compared to problem behaviors and the affects of gender on informants' ratings. Exploration of these factors informs researchers and clinicians to develop best methods for incorporating parent and teacher reports when using the BASC-2 in assessing children and adolescents.

CHAPTER 3

METHODS AND PROCEDURES

Demographic Information

The evaluations that were selected for review in this study included children and

adolescents ages 6-17 (M=10.81, SD=3.177) who were referred and evaluated for a

learning disorder or socio-emotional disorder in a university based outpatient clinic. In

regards to gender, 34.7 percent were identified as female and 65.3 as male. 3.4 percent of

the sample were identified as African American, 1.1 percent as Asian American, 1.1

percent as Native American, 92 percent of the sample were identified as White, and 2.3

percent of the sample as Biracial. Over half of the children (57.9 percent) were in grades

K-5. Type of family was coded for Intact (two biological parents, 71.3 percent); Blended

(Step-families, 16 percent); Single Parents (3.2 percent); Other (which included

grandparents, adoptive parents and same sex parents, 9.6 percent). The evaluations were

selected for the study if BASC-2 were completed as part of the assessment. All

demographic information has been included below in Table 3.1 Inclusion criteria for the

study included those: (a) have been diagnosed with an internalizing disorder (Mood

Disorder NOS, Depressive Disorder NOS, Major Depression, Dysthymia, Anxiety

Disorder NOS and Generalized Anxiety Disorders); (b) have been diagnosed with an

externalizing disorder (Conduct and Oppositional Defiant Disorders); (c) diagnosis of

ADHD (Inattention, Hyperactive/Impulsive, and Combined Type); (d) other – Learning

Disorders, Pervasive Developmental Disorders and Reactive Attachment Disorders and

(e) have no diagnosis.

Table 3.1

Descriptive Statistics for Demographic Variables*

	Mean	Standard Deviation
Age	10.81	3.77
	Frequency	**Percent**
Gender		
Female	34	34.7
Male	64	65.3
Missing	2	
Race		
African American	3	3.4
Asian American	1	1.1
Native American	1	1.1
White/Caucasian	81	92
Biracial	2	2.3
Missing	12	
Grade		
Elementary (K-5)	55	57.9
Middle (6-8)	17	18
High School (9-12)	23	24.1
Missing	5	
Type of Family**		
Blended	15	16
Intact	67	71.3
Single	3	3.2
Other	9	9.6
Missing	6	
Diagnosis		
Internalizing Disorders	17	17.7
Externalizing Disorders	2	2.1
Attention Disorders	20	20.8
Other Disorders	29	30.2
No Diagnosis	28	28.0
Missing	4	

Number of Siblings**		
Only Child	18	19.4
1 sibling	28	30.1
2 siblings	30	32.3
3 siblings	14	15.1
4 siblings	3	3.2
Missing	7	

Note. * Includes original data set (100 evaluations) with missing values **Gathered as additional information, but not used in analyzing the four research questions.

Description of the Instrument

Behavioral Assessment System for Children, Second Edition (BASC-2) is a multi-method assessment tool used to assess adaptive skills and problem behaviors in children and adolescents (Reynolds & Kamphaus, 2004). The system provides multiple rater forms including self-reports, parent and teacher reports. Parent and teacher reports will be examined for the purposes of this study. The forms include a 4-point likert scale for behaviors observed: 0 = Never, 1 = Sometimes, 2 = Often, 3 = Almost Always.

The BASC-2 can be hand or computer scored (BASC-ASSIST) to obtain profiles with T-scores of all the scales. Scored profiles will be used to collect data for the purposes of this study. A brief description of the scales used in this study is found in Appendix B and the BASC-2 scale structure and scoring system can be found in Appendix C.

Psychometric Properties

The psychometric properties of the BASC-2 (Reynolds & Kamphaus, 2004) have been updated since its original version. The reliability and validity of the BASC-2 range from acceptable to excellent. The internal consistency of each individual scales is between .80 and .90 across teacher and parent forms. The composite scales are between

.90 and .97. Test-retest reliability based on testing from one to eight weeks after the initial administration were between .80 and .90 for composite scores and for individual scales between .60 and .90 across parent and teacher forms for all age groups. Interrater reliability was computed for each form and across parent and teacher reports. The interrater reliability between teacher forms was between .48 and .81 and for the parent forms was between .65 and .85 for composite scores. For each subscale the interrater reliability ranged from .19 to .82 for the teacher and parent forms. Correlations across parent and teacher reports for composite scores ranged from .22 to .51. Interrater reliability for the Internalizing subscale was lower than the other composite scores.

The validity of the BASC-2 has been tested across content, criterion and construct validity. Content validity refers to the extent that the items on the BASC-2 are able to represent the construct being measured. BASC-2 was created utilizing information from multiple sources including teachers, parents, children and psychologists. The scales were also developed using DSM-IV-TR (2000) criteria. Criterion validity is the ability to which the BASC-2 meets criteria that has already been established. The clinical scales have been shown to be consistent with the DSM-IV-TR (2000) criteria. Also, on the clinical scales where demographic scales were controlled for, the Internalizing and Externalizing scale have been effective in distinguishing between clinical and non-clinical samples (Reynolds & Kamphaus, 2004). Construct validity is the extent to which the constructs measure what it purports to measure. The BASC-2 has shown to be correlated with the Parent and Teacher reports of the CBCL (Achenbach et.al, 2008; Doyle, Ostrander, Skare, Crosby, & August, 1997) and the original BASC (Waggoner,

2005). Both the teacher and parent forms have been shown to be effective in identifying children with and without ADHD (Lett & Kamphaus, 1997; Jarratt, Riccio, & Siekierski, 2005, Doyle, Ostrander, Skare, Crosby, & August, 1997). Regarding external validity the BASC-2 was normed based on current U.S. Census population (Reynolds & Kamphaus, 2004). There is limited research of applicability across various cultures; more research needs to be conducted on persons of color, immigrant populations, across various socio-economic levels, and other multicultural factors. Reliability as measured by internal consistency, test-retest and interrater reliability has been found to be good. Validity across content, construct and criterion has been found to be moderate to high (Frick, Barry, & Kamphaus, 2009). Research to date on the BASC-2 has supported its sound psychometric properties.

Procedures

A chart review study was conducted using records from a university based outpatient clinic that provided assessment and therapy services for a small southeastern community in the United States. The evaluations were conducted on children and adolescents from January of 2004 to January of 2011. Only charts that used the BASC- 2 (Reynolds & Kamphaus, 2004) teacher and parent rating scales were reviewed for this study. A typical evaluation included 12-15 assessment measures used in conjunction with clinical interview and observation of the child. All clients were clinically evaluated by doctoral-level Counseling Psychology students under the supervision of a licensed psychologist and diagnoses assigned based on clinical interview with parents and children, observation of the child, intellectual and achievement testing (Weschlers and

39

Woodcock Johnson tests), behavior rating scales (BASC-2 or CBCL), self-report scales (Child Depression Inventory, Revised Child Manifest Anxiety Scale etc.), and projective tests (Sentence completion). These evaluations included two days of testing and additional tests were added on the second day based on information gathered from clinical interview and observations. The assessment data was then reviewed during group supervision and individual supervision by two licensed psychologists (one with a specialization in Learning Disabilities and the other in Socio-Emotional Disorders). Demographic variables were gathered from these evaluations including: age of child, race of child, gender of child, household composition [Type of family (intact family, blended family, single family, or other); Number of siblings], grade of child, and subject taught by teachers (academic or non-academic).

The researcher gathered T-scores from profiles on the scales related to adaptive skills, internalizing and externalizing behaviors, overall behavior problems and attentional problems for the parent and teacher reports. More specifically, the BASC-2 includes Aggression, Conduct problems and Hyperactivity subscales under the Externalizing scale. On the Internalizing scale, the BASC-2 is comprised of Anxiety, Depression and Somatization subscales. The BASC-2 also includes a separate Attention Problems subscale. The Behavioral Symptom Index is an overall measure of problem behaviors that includes Aggression, Hyperactivity, Anxiety, Depression, Attention Problems and Atypicality subscales. The Adaptive Skills scale includes Activities of Daily Living (for parent form), Study Skills (for teacher form); and for both forms Leadership, Social Skills and Functional Communication subscales. Lastly, the diagnosis

40

based on the DSM-IV-TR (2000) for the child will be gathered including evaluations that reported no diagnosis. A sample chart-review form is provided in Appendix A.

Hypotheses

Based on the current research and findings the researcher predicts the following for each research question:

Research Question 1

What is the level of agreement between ratings of different informants (mothers, fathers, and teachers) on the BASC-2 for each of the informants on the five subscales?

(a) Mothers and Fathers will have higher levels of agreement on all scales except Internalizing.

(b) All teachers will have higher levels of agreement on all scales except on the Internalizing Scale.

(c) Across parents and teachers, raters will have higher levels of agreement on Externalizing, Behavioral Symptom Index and Attention Problems scales.

(d) Across parents and teachers, raters will have lower levels of agreement on Internalizing and Adaptive skills scales.

Research Question 2

How do different informants differentially and/or similarly predict the final diagnosis (including having no diagnosis)?

(a) Parents' scores will more often predict the final diagnosis compared to teachers for internalizing diagnoses.

(b) Parents' scores will be sufficient to predict the diagnosis of an internalizing disorder.

(c) Teachers' scores will more often predict the final diagnosis compared to parents' scores on externalizing and attention related diagnoses.

(d) Teachers' scores will be sufficient to predict the diagnosis of an externalizing disorder and in the diagnosis of attention problems.

(e) Parents' and Teachers' scores will be needed to predict no diagnosis or "other" diagnoses (i.e. Learning Disorders, etc.).

Research Question 3

Do informants' perceptions differ on positive behavior compared to problem behavior?

(a) Parents will report more positive behavior as measured by the Adaptive Skills scale than teachers.

(b) Teachers will report more problem behaviors as measured by the Behavioral Symptom Index.

Research Question 4

Does gender of the child interact with the informants ratings?

(a) There will be gender differences in informant ratings based on the gender on the Behavioral Symptom Index.

(b) Boys will have more problem behaviors on the Externalizing and Attention Problems scales compared to girls.

(c) Girls will have more problems on the Internalizing Scale compared to boys.

Statistical Analysis

Missing Values Treatments

Through power analysis for a regression analysis with five predictors it was determined that at least 91 evaluations were needed for a medium effect size (.15) (Soper, 2011). Although 100 evaluations were included in this study, the data included missing values across a number of variables limiting the analysis. In reviewing literature on missing values, there are multiple methods to address missing values with advantages and disadvantages to each one. A brief description of each method is listed below (Scheffer, 2002):

Listwise deletion or complete case analysis is a method used by deleting cases in which any data points are missing. Another similar method includes deleting any variables that may have concentrated missing data. Pairwise deletion only removes cases that have missing values involved in the specific analyses such as correlations. The advantage to this is that you have clean data without including any estimated data. However, this only works when loss of power is tolerable, percentage missing is small, and is concentrated on a few variables. The disadvantage is that if the points that are deleted vary considerably from other data points this could bias the data. Additionally, significant portions of the data may be lost which only allows for simple data analysis.

Single imputation methods are also used with simple or complex strategies such as mean substitution, regression substitutions and hot deck imputation. A common and easy method that is often used is imputing means. The advantage to this imputation method is that it is easy to impute and depending on the pattern of missing data can solve

the problem. It works well when small percentage of data is missing. However, a major downside is that it can give a biased estimate of the variance and covariance. This is especially an issue when more complex analyses are considered and p-values are involved. There are other more sophisticated methods for single imputation such as Expectation Maximization (EM); however, require that no more than 5 to 10 percent of the data are missing.

In a recent article, the *Best Practices for Missing Data Management in Counseling Psychology* (Scholmer, Bauman, & Card, 2010) was discussed. The authors reviewed three methods of handling missing data: mean substation, multiple imputation and full information maximum likelihood. Multiple imputation and full information maximum likelihood were seen as preferred alternatives to deletion or single imputation methods. The authors urge counseling psychology as a field to move towards reporting and addressing missing data based on best practices.

Multiple Imputation creates multiple sets of data using different algorithms (i.e. EM, stochastic regression, and Markov Chain Monte Carlo) to impute data so that there is variation across the data sets. Scholmer, Bauman & Card (2010) stated in their article, "the precision of parameter estimated and accuracy of standard errors make MI one of the best options for handling missing data (p.5)." The different data sets (usually 3-5) are analyzed separately and then the results of each data set are pooled together to obtain a single result. The advantages of multiple imputation are that unbiased estimates can be derived, there is no loss of power and works with any analysis. It is also reliable even with up to 50 percent of missing data. The disadvantages are ensuring the best imputation

model is used which can be time consuming. Furthermore, not all software packages include pooled results for all analyses (Scholmer, Bauman & Card, 2010).

Missing Values Patterns

The current data set was analyzed to determine the patterns of missing data which are presented below. The table below provides information on the original data set and includes the frequency and missing values across variables. As can be seen in Table 2.2 there are 55 to 95 values missing out of a 100 across the subscales for Teachers 3, 4, & 5. Therefore data set was only used for two teachers. As such for evaluations that included more than two teachers, two teachers were then selected at random from the available three, four or five teachers.

Table 3.2
Frequency and missing values for variables in original data set

	Frequency of data N							
	Valid	Missing		Valid	Missing		Valid	Missing
DOB*	100	0	TI1	89	11	TI5	5	95
Age*	100	0	TE1	90	10	TE5	5	95
Gender	98	2	TAP1	90	10	TAP5	5	95
Race*	88	12	TBSI1	89	11	TBSI5	5	95
Grade*	95	5	TA2	72	28	PA1	94	6
Family*	94	6	TI2	73	27	PI1	95	5
Siblings*	93	7	TE2	73	27	PE1	95	5
Birth order*	92	8	TAP2	73	27	PAP1	95	5
Diagnosis	96	4	TBSI2	73	27	PBSI1	94	6
TLTK1*	82	18	TA3	47	53	PA2	80	20
TLTK2*	64	36	TI3	45	55	PI2	80	20
TLTK3*	40	60	TE3	47	53	PE2	80	20
TLTK4*	13	87	TAP3	47	53	PAP2	80	20
TLTK5*	5	95	TBSI3	45	55	PBSI2	80	20
TAS1*	86	14	TA4	15	85			

TAS2*	71	29	TI4	15	85
TAS3*	46	54	TE4	15	85
TAS4*	15	85	TAP4	15	85
TAS5*	4	96	TBSI4	15	85
TA1	88	12	TA5	7	93

Note. T = Teacher and P = Parent; P1 = Mother and P2 = Father; Family = Type of family (Intact, Blended, Single parent etc.)TLTK (Length of Time Known for Teacher 1, 2, 3, 4, 5); TAS (Subject taught by teacher – Academic/Non Academic); A = Adaptive Skills; I = Internalizing; E = Externalizing; AP = Attention Problems; BSI = Behavior Symptom Index; *Gathered as part of data collection and demographic information, but not included as variables in the research questions.

Since values were still missing across most variables, a missing values analysis was conducted on the data set including two teachers. Only variables directly included in the specific research questions were included in the missing values analysis. These variables included, Gender, Diagnosis, five subscales for Teachers 1 and 2 and subscales for Mothers and Fathers. As seen below in Chart 1, if listwise deletion is used 47 percent of the cases would be deleted with only 53 percent of the data points available for analysis. Additionally, 14.68 percent of the data points are missing across all variables, therefore single imputation methods would not be appropriate. Based on the pattern of missing data which was found across parent and teacher variables and Little's chi-square statistic (Little, 1989) for testing whether values are missing completely at random (MCAR); the pattern of data in this study are missing at random (MAR); therefore, a multiple imputation method is most appropriate for this data set. According to Graham & Schafer (1999), the following conclusions were made based on their simulation study, "missing data only make the problem worse leading to an effective sample size even smaller than before...the good news, however, is that provided one has sufficiently strong effects, multiple imputation...can be an excellent solution to the missing-data problem,

46

even with sample sizes as low as 50, even with as much as 50% missing from most

variables, and even with relatively large and complex models (p.29)."

Figure 3.1

Missing values summary

Overall Summary of Missing Values

Variables	**Cases**	**Values**
	Panel	
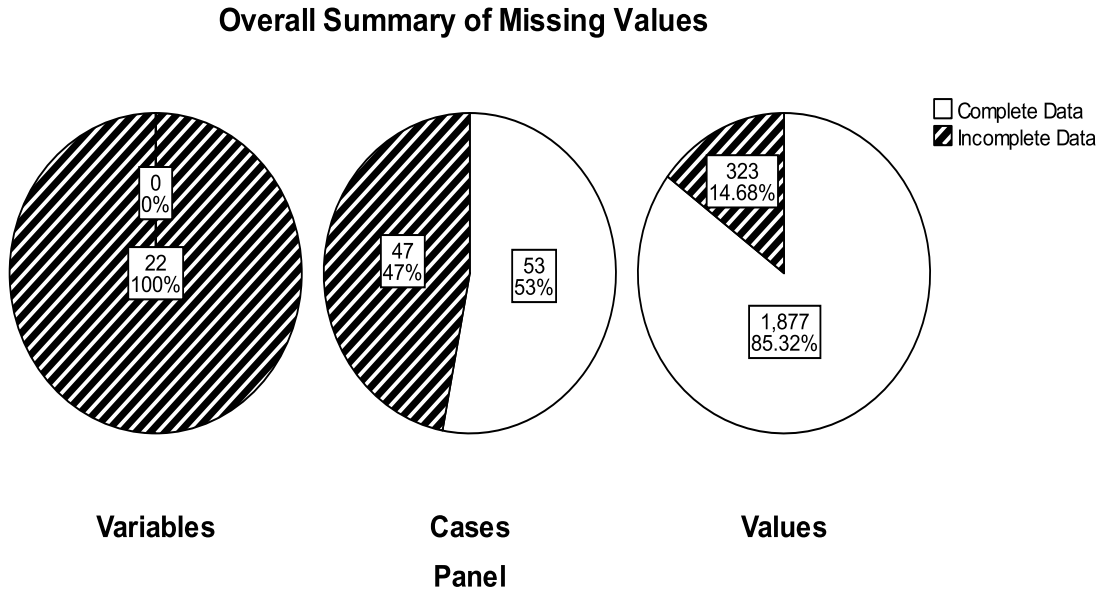
Multiple imputation methods were computed using SPSS (version 19). Five

separate data sets were created with imputed data for missing values. A linear regression

model was used for the variables to impute missing values. The following Table 3.2

shows an example of a variable (PE2) and the means and standard deviations between

each data set. As seen below 20 data points were imputed for this variable. The mean for

original data was 54.21 and the mean for the imputed values range from 54.41 to 55.71.

The data analysis for Research Questions 1 and 2 in this study used the five imputed data

sets to obtain pooled results. Research Questions 3 and 4 used the original data set for

analysis as variables were collapsed and SPSS (version 19) does not provide pooled

results for these statistics.

Table 3.3

Comparison of imputed data and original data (Means and Standard Deviations)

Data	Imputation	N	Mean	Std. Deviation	Minimum	Maximum
PE2 – Father Externalizing subscale						
Original Data		80	54.21	10.519	36.00	80.00
Imputed Values	1	20	57.54	13.986	33.61	83.07
	2	20	61.69	11.827	41.97	84.16
	3	20	56.38	12.260	32.52	80.14
	4	20	55.21	8.590	41.38	72.01
	5	20	61.50	7.369	44.17	73.11
Complete Data After	1	100	54.88	11.297	33.61	83.07
Imputation	2	100	55.71	11.143	36.00	84.16
	3	100	54.65	10.858	32.52	80.14
	4	100	54.41	10.130	36.00	80.00
	5	100	55.67	10.359	36.00	80.00

Summary of Statistical Analysis by Research Question

Research Question 1 examined the level of agreement between mothers, fathers

and two teachers. To test the hypotheses for this question, a Pearson-Product Moment

correlation was used with the five imputed data sets. Specifically, a correlation was

completed between the five subscales (Internalizing, Externalizing, Attention Problems,

Behavior Symptom Index, and Adaptive Skills) and the four informants (mother, father,

teacher 1 and teacher 2). Additionally, to test the correlation between parents and

teachers as two groups, an Interclass Correlation was calculated (ICC).

Research Question 2 examined the ability of informants' to predict the final

diagnosis. To test the hypotheses for this question, a stepwise logistic regression analysis

(Entry: Backward-Likelihood) was conducted with the five imputed data sets. The data

sets were stacked and weighted for analysis. Specifically, each diagnosis was categorized into five categories (internalizing, externalizing, attention disorders, no diagnosis and other). A separate regression analysis was completed for each diagnostic category with all four informants included in each analysis. In the "no diagnosis" category, Adaptive Skills for each informant was used as the independent variable. In the "other" category, the Behavior Symptom Index for each informant was used as the independent variable.

Research Question 3 compared the difference between parents and teachers ratings on positive behavior and problem behavior using the original data set. To test the hypotheses for this question means were calculated for parents and teachers for Adaptive Skills and Behavior Symptom Index. Then two separate independent t-tests were run to compare the differences in the means.

Research Question 4 focused on the interaction of the child's gender on informant's ratings. The hypotheses were tested using Multivariate Analysis of Variance (MANOVA) on the original data set. Multivariate analysis of variance was conducted to determine whether there were significant gender differences for each scale across parent and teacher ratings. Box's M test and descriptive tests were used to determine which gender had differences for teacher and parents.

CHAPTER 4

RESULTS

Results will be presented using multiple imputation data sets to derive a pooled

result for research questions 1 and 2. The results will be presented in the following order:

descriptive statistics and tests of hypotheses for each research question. These tests

include correlations (Pearson Product Moment and Interclass Correlations); independent

t-tests, logistic regression analysis; and Multivariate Analysis of Variance (MANOVA).

Descriptive Information

Data was collected for 100 evaluations (that met criteria for inclusion in the

study) out of 471 evaluations completed from January 2000 to May 2011 from a

university based outpatient clinic that provides assessment and therapy services. These

evaluations were conducted by doctoral level counseling psychology students and

supervised by a team of licensed psychologists. For this study two parents and two

teachers were included as informants on the BASC-2. When more than two teachers were

available, random selection of teachers was used to include in the study. Table 4.1

includes descriptive information for each informant under study with mean *t* scores (*SD*

were not calculated for pooled results) for pooled data. To confirm normal distributions

of the subscales, measures of skewness and kurtosis were computed for each informant

across subscales. Skewness and kurtosis values of zero are indicative of a normal

distribution, and values between -2 and +2 signify no problematic deviations from

normality (Bulmer, 1979, p 64). All measures of skewness and kurtosis for all scales

were between the values of -1.5 to +1.5 across all data sets. Pooled results were not computed for skewness and kurtosis.

Table 4.1

Pooled Means for Variables in Study

Pooled Results	Mean		Mean
Teacher 1 Adaptive Skills	43.82	Mother Attention Problems	57.75
Teacher 1 Internalizing	56.40	Mother Behavior Symptom	56.49
Teacher 1 Externalizing	54.30	Father Adaptive Skills	43.36
Teacher 1 Attention Problems	57.20	Father Internalizing	55.07
Teacher 1 Behavior Symptom	55.85	Father Externalizing	54.33
Teacher 2 Adaptive Skills	45.81	Parent Attention Problems	57.24
Teacher 2 Internalizing	53.49	Parent Behavior Symptom	55.51
Teacher 2 Externalizing	51.42		
Teacher 2 Attention Problems	55.99		
Teacher 2 Behavior Symptom	53.23		
Mother Adaptive Skills	43.70		
Mother Internalizing	56.38		
Mother Externalizing	54.61		

Note. $N = 100$ for all variables

Hypothesis Testing

Data analyses were conducted for each research question and hypotheses using SPSS (version 19). The significance level was set at .05. Each question and hypotheses (found in Chapter 3) are addressed below with corresponding analyses.

Level of Agreement (Research Question 1)

What is the level of agreement between ratings of different informants (mothers, fathers, and teachers) on the BASC-2 for each of the informants on the five subscales?

Pearson-Product Moment Correlations (Table 4.2) were calculated for Teacher (teacher 1 and teacher 2) and Parent (mother and father) variables across the 5 subscales. Correlations for all informants within the same setting (home/school) across subscales will be presented and followed by comparisons between parents and teachers.

Across the Internalizing scale, the ratings of mothers and fathers had a strong correlation, $r(100) = .621$ $p = .000$. Ratings of teacher 1 and teacher 2 had a moderate correlation $r(100) = .391$ $p = .000$ On the Externalizing scale, the ratings of teachers (1 and 2) had a strong correlation, $r(100) = .625$ $p = .000$ with similar findings for ratings of mothers and fathers, $r(100) = .560$ $p = .000$. Likewise, on the Attention Problems scale ratings of mothers and fathers had a strong correlation $r(100) = .602$ $p = .000$ and ratings of teachers (1 and 2) had a strong correlation $r(100) = .500$ $p = .000$. For Adaptive Skills mothers and fathers had a strong relationship $r(100) = .634$ $p = .000$. However, teachers (1 and 2) had a moderate relationship $r(100) = .419$ $p = .000$ on the Adaptive Skills scale. Mothers and fathers had a strong relationship for Behavior Symptom Index, $r(100) = .597$ $p = .000$ as did both teachers, $r(100) = .591$ $p = .000$.

52

Across parents' and teachers' ratings the correlations ranged from having no

significant relationship to having a moderate relationship. On the Internalizing subscale

the ratings between parents and teachers ranged from a weak to moderate correlation for

teacher 1 and having no significant relationship between ratings of mothers and teacher 2.

On the Externalizing subscale ratings across parents and teachers had a moderate

relationship. Parents' and teachers' ratings had a weak to moderate correlation for the

Attention Problems subscale. Parents' and teachers' ratings had no significant

relationship to a weak relationship for the Adaptive Skills subscale. Lastly, on the

Behavior Symptom Index (BSI) subscale weak to moderate relationships were found for

ratings across all teachers and parents.

Table 4.2

Pooled Pearson-Product Moment Correlations across Subscales

Scales	Mothers/ Fathers	Teacher 1/ Teacher2	Mothers/ Teacher 1	Mothers/ Teacher 2	Fathers/ Teacher 1	Fathers/ Teacher 2
Internalizing	.621**	.391**	.337**	.221	.281*	.382*
Externalizing	.560**	.625**	.411**	.370**	.340*	.248*
Attention Problems	.602**	.500**	.332**	.340*	.359*	.282*
Adaptive Skills	.634**	.419**	.225*	.158	.258*	.309*
Behavioral Symptom Index	.597**	.591**	.410**	.405**	.275*	.263*

Note. *p < .05, ** p < .01

The Pearson-Product Moment correlations were then compared for teachers and parents using Fisher's *r*-to-*z* transformation to determine significant differences between the correlations. A *z* score of 1.96 or above indicates significance difference. Significant difference in correlations were found for ratings of teacher 1 and teacher 2 for the Internalizing and Externalizing scales ($z = 2.23$) and Externalizing and Adaptive Skills ($z = 2.00$). No significant differences were found for mothers and fathers across subscales.

Correlations examine the extent of relationships; however, to test the level of agreement an Interclass Correlation (ICC) was calculated. Fleiss (1986) includes benchmarks to interpret the ICCs : > .75 = excellent; .40 to .75 = fair to good; and .40 = poor. As seen in Table 4.5 there are two measures that are reported: single measures and average measures. The single measure ICC gives you the reliability for a single informant's rating, that is, whether the ratings for one informant are similar to that of another. The average measure ICC gives you the reliability of the mean of all ratings for all raters. This indicates whether the ratings are stable when the ratings are averaged across all raters. As SPSS (version 19) does not calculate a pooled ICC. Therefore based on Schafer & Graham (2002), ICCs were derived for each data set and then a combined average was calculated. Separate ICC correlations were calculated for each scale for mother and fathers as a group and then teacher 1 and teacher 2 as another group. A Fisher's *r* to *z* transformation was calculated to compare the two correlations. The Table 4.3 presents this information and includes pooled single measures ICC for each scale. A significant difference was found for Adaptive Skills ($z = 2.2$) and Internalizing subscales ($z = 2.0$) for parents and teachers.

Table 4.3

Pooled Interclass Correlations for Parents and Teachers

Subscales	Parents	Teachers	*r to z*
Internalizing	.610	.373	2.21*
Externalizing	.560	.588	.29
Attention Problems	.602	.495	1.07
Behavior Symptom Index	.587	.561	.28
Adaptive Skills	.631	.404	2.2*

Note. Absolute value of *z* is reported. **z >1.96*

Summary of Research Question 1

Based on the Pearson-Product Moment Correlations and ICCs some of the proposed hypotheses were supported. Although it was hypothesized that the agreement for ratings of mothers and fathers would be lower for the Internalizing scale, this was not the case. The ratings of mothers and fathers in general had high levels of agreement/relationship across all the scales. Ratings of teachers (1 and 2) had moderate levels of agreement/relationship across all scales except the Externalizing scale. On this scale, teacher 1 and teacher 2 ratings had a strong relationship. There was also a significant difference between the two teachers' ratings on the Internalizing and the Externalizing scale. Comparing the level of agreement between parents' and teachers' ratings across subscales, there was a significant difference for Internalizing and Adaptive Skills which supported the hypothesis that agreement between ratings of parents and teachers would be higher for Externalizing, Behavioral Symptom Index, and Attention

55

Problems scales. The ratings of parents and teachers had lower levels of agreement on Internalizing and Adaptive Skills. Overall, the findings suggest that there are higher correlations and stronger agreement among informants within the same setting (home/school) which was expected. Additionally, higher levels of agreement was found for ratings between parents and teachers for externalizing behaviors and lower levels of agreement for ratings between parents and teachers for internalizing and positive behavior. Effect sizes were calculated for the correlations and ranged from $R^2 = .043$ to .659 (for significant correlations).

Predictive Ability (Research Question 2)

How do different informants differentially and/or similarly predict the final diagnosis (including having no diagnosis)?

Stepwise logistic regression analysis using backward entry was conducted to determine whether teacher and parent ratings on each scale predicted the category of diagnoses (internalizing, externalizing, attention disorders, no diagnosis and other diagnoses). As SPSS does not calculate a pooled estimate of logistic regression, another method for pooling results using the five data sets was calculated. According to Wood, White & Royston (2008) the best method for pooling results for a logistic regression when using variable selection is computing estimates based on Rubin's Rules (Rubin, 1987). However, this method can be computed when calculating Wald statistics, but not for Odds Ratio (Likelihood Ratio tests). Likelihood Ratio tests are the preferred statistic compared to the Wald tests especially given small sample sizes (Fox, 1997). Given this, stacked imputed data sets with weighted regression were used (Wood, White & Royston,

2008). As described in the article, the following formula was used to weight the data:

W2: $wi = (1-f)/M$ where M is the number of imputed data sets and f is the average

fraction of missing data across all variables, that is calculated as (total number of missing

values across all variables) divided by pn (number of cases/participants). The five data

sets were stacked and variables were weighted ($wi=.147$) and a stepwise logistic

regression (backward entry) was calculated. The results for each regression are presented

below.

Stepwise logistic regression analysis using backward entry was conducted to

determine whether teacher and parent ratings on the Internalizing scale could predict an

internalizing diagnosis. Regression results indicate that the overall model for step 2

which included (mothers, fathers and teacher1) was significant, $\chi^2(3) = 8.111$, $p =.044$, R^2

$= .104$. The model correctly classified 79.2 percent of the cases in terms of the presence

or absence of an internalizing diagnosis; however standardized regression coefficients

were not significant. The overall model for step 3 (which included teacher 1 and fathers)

was also significant $\chi^2(2) = 7.327$, $p =.026$, $R^2 = .095$. The model classified 79.8 percent

of the cases in terms of a presence or absence of an internalizing disorder. The last model

(step 4, included only fathers) was significant, $\chi^2(1) = 6.753$, $p =.009$, $R^2 = .88$. The

model classified 80.4 percent of the cases in terms of a presence or absence of an

internalizing disorder. Examination of standardized regression coefficients for step 3 and

4 revealed that fathers' rating (Odds Ratio [OR] = 1.08) was a significant predictor in the

overall model, such that those with higher scores on the Internalizing scale (on the Parent

Report) were 1.08 times as likely to have an internalizing diagnosis than not (see Table

4.4).

Table 4.4

Stepwise Regression Models (Backward Entry) for Predicting Internalizing Disorders

from BASC-2 Internalizing Scale

		B	S.E.	Exp(B)	Cox & Snell R^2
Step 1	TI1	-.021	.028	.979	.105
	TI2	-.004	.031	.996	
	PI1	.025	.029	1.025	
	PI2	.061	.036	1.063	
	Constant	-5.184	2.157	.006	
Step 2	TI1	-.022	.025	.978	.104
	PI1	.026	.029	1.026	
	PI2	.059	.034	1.061	
	Constant	-5.271	2.057	.005	
Step 3	TI1	-.018	.025	.982	.095
	PI2	.075	.030	1.078*	
	Constant	-4.885	1.961	.008	
Step 4	PI2	.069	.028	1.071*	.088
	Constant	-5.539	1.747	.004	

Note. p<.05; TI1=Teacher 1 Internalizing Scale; TI2 = Teacher 2 Internalizing Scale; P1=
Mother Internalizing Scale; P2 = Father Internalizing Scale

Stepwise Regression results to determine whether teacher and parent ratings on

Attention Problems could predict attention-related diagnosis indicate that the overall

model for step 1 (all four informants) was significant, $\chi^2(4) = 10.232$, $p =.037$, $R^2 = .130$.

The model correctly classified 78.0 percent of the cases in terms of the presence or

absence of attention-related diagnosis. Examination of standardized regression

coefficients revealed that none of the predictors were significant in the model. For step 2

(teacher1, teacher2, and mothers), the overall model was significant, $\chi^2(3) = 9.485$, p

$=.023$, $R^2 = .121$. The model correctly classified 77.4 percent of the cases in terms of the

presence or absence of attention-related diagnosis. Parent 1(mothers) rating (OR = 1.07)

was a significant predictor in the overall model, such that those with higher scores on the

mother-rated Attention Problems scale were 1.07 times as likely to have an attention-

related diagnosis than not (see Table 4.5).

Table 4.5

Stepwise Regression Models (Backward Entry) for Predicting Attention Disorders from

BASC-2 Attention Problems Scale

		B	S.E.	Exp(B)	Cox & Snell R^2
Step 1	TAP1	.051	.033	1.052	.130
	TAP2	-.058	.033	.944	
	PAP1	.058	.042	1.059	
	PAP2	.033	.038	1.033	
	Constant	-6.540	2.564	.001	
Step 2	TAP1	.055	.032	1.057	.121
	TAP2	-.056	.033	.946	
	PAP1	.073	.037	1.076*	
	Constant	-5.906	2.425	.003	

Note.*p<.05; TAP1= Teacher 1 Attention Problems Scale; TAP2 = Teacher 2 Attention
Problems Scale; PAP1= Mother Attention Problems Scale; PAP2 = Father Attention
Problems Scale

Regression results for predicting "No Diagnosis" using Adaptive Skills and

"Other" using BSI indicate that the overall model was not significant for predicting either

category. Data analysis that was planned for the externalizing disorders was not

conducted due to insufficient data on the externalizing diagnosis category.

Summary Research Question 2

Overall, results indicate that fathers' ratings on the Internalizing scale were

predictive of an internalizing diagnosis. Additionally on the Internalizing scale, though

individual predictors were not significant, overall models were significant and included

ratings of mothers, fathers, and teacher1. Mothers' ratings on the Attention Problems

scale were predictive for attention-related diagnosis. Furthermore, overall models for

Attention Problems scale included all four of the informants' ratings with only teachers'

and mothers' ratings significant in the next step of the model. The hypotheses that were

supported by these results include that parents' ratings were sufficient to predict

diagnosis of an internalizing disorder. Although, teachers' ratings were significant in

overall models, the hypothesis that teachers' ratings would be sufficient for externalizing

disorders was not supported. Lastly, neither teachers' nor parents' ratings predicted the

"No Diagnosis" category or "Other" category which refuted the hypothesis that all four

informants were needed for predicting these categories of diagnosis.

Positive vs. Problem Behavior (Research Question 3)

Do informants' perceptions differ on positive behavior compared to problem behavior?

Parents and teachers were grouped across Adaptive Skills (AS) and Behavior

Symptom Index (BSI) scales to compare their ratings (N=287). Therefore, the original

data set was used for analysis. Independent samples *t* test was run to determine whether

there were differences in Adaptive Skills and Behavior Symptom Index ratings across

parents and teachers. Levene's test for homogeneity of variance was not significant,

indicating that this assumption was not violated. Results indicated that there were no

significant differences, $t(332) = .961$, $p = .802$, $d = .10$. in Adaptive Skills ratings across

parents and teachers nor were there differences in BSI ratings, $t(333) = -.758$, $p = .815$, d

$= .08$. (see Table 4.6). The proposed hypotheses that, (a) Parents will rate children higher

60

on adaptive behavior as measured by the AS scale than teachers (b) Teachers will report

more problem behaviors as measured by the BSI were not supported.

Table 4.6

Independent Samples t-test for Parents and Teachers for Adaptive Skills and Behavior Symptom Index Scales

	Group	*N*	*Mean*	*SD*	*t*	*df*	*sig.*
Adaptive Skills	Teachers	160	44.38	8.975	.967	332	.334
	Parents	174	43.40	9.373			
Behavior Symptom Index	Teachers	162	55.17	11.893	-.758	333	.449
	Parents	173	56.14	11.437			

Gender Interaction (Research Question 4)

Does gender of the child interact with the informants ratings?

Multivariate analysis of variance was conducted to determine whether there were

significant gender differences for the five subscales across parent and teacher ratings.

Each subscale was analyzed separately. Box's test for equality of covariances was

computed. Additionally, Pillai's Trace was used as the multivariate statistic, as this is the

most robust of the four leading tests of group differences (Olson, 1976) especially if

assumptions are violated (i.e.when sample size decreases, unequal cell sizes or

homogeneity of covariances is violated).

For the Internalizing subscale across parent and teacher ratings, the Box's test for

equality of covariances revealed that the assumption of homogeneity of covariance was

not violated. However, multivariate analyses indicated no significant differences across

genders on the multivariate combination of the parent and teacher ratings (see Table 4.7).

On the Externalizing subscale across parent and teacher ratings, the Box's test for equality of covariances revealed that the assumption of homogeneity of covariance was violated. Multivariate analyses indicated significant differences across genders on the multivariate combination of the parent and teacher ratings, Pillai's Trace = .010, $F(2, 139) = 7.58$, $p = .001$, partial $\eta^2 = .09$. Univariate analyses revealed significant differences between the groups on [Teachers], $F(1, 140) = 11.11$, $p = .001$, partial $\eta^2 = .07$, and [Parents], $F(1, 140) = 9.96$, $p = .002$, partial $\eta^2 = .67$. Post hoc tests were not performed for Gender because there are fewer than three groups. Comparing the means for Parents' rating for boys ($M = 55.54$, $SD = 11.12$) and girls ($M = 49.44$, $SD = 8.98$) indicate that Parents rated boys higher on the Externalizing scale. In comparing the means for Teachers' ratings on the Externalizing scale, results indicate that boys ($M = 56.13$, $SD = 11.90$) were rated higher than girls ($M = 50.26$, $SD = 7.60$) (see Table 4.7).

On the Attention Problems subscale across parent and teacher ratings, the Box's test for equality of covariances revealed that the assumption of homogeneity of covariance was not violated. Multivariate analyses indicated significant differences across genders on the multivariate combination of the parent and teacher ratings, Pillai's Trace = .49, $F(2, 139) = 3.59$, $p = .03$, partial $\eta^2 = .05$. Univariate analyses revealed significant differences between the groups on [Teachers], $F(1, 140) = 6.084$, $p = .15$, partial $\eta^2 = .042$, but not on [Parents], $F(1, 140) = 3.372$, $p = .07$, partial $\eta^2 = .02$. Post hoc tests were not performed for Gender because there are fewer than three groups. Teachers rated boys ($M = 59.17$, $SD = 9.98$) higher than girls ($M = 54.70$, $SD = 10.93$) for Attention Problems scale (see Table 4.7).

The Box's test for equality of covariances on the Adaptive Skills subscale for parents and teachers revealed that the assumption of homogeneity of covariance was not violated. Multivariate analyses indicated significant differences across genders on the multivariate combination of the parent and teacher ratings, Pillai's Trace = .08, $F(2,136)$ = 5.810, p = .004, partial η^2 = .10. Univariate analyses revealed significant differences between the groups on [Teachers], $F(1, 137)$ = 9.90, p = .002, partial η^2 = .067, and [Parents], $F(1, 137)$ = 4.265, p = .04, partial η^2 = .03. Post hoc tests were not performed for Gender because there are fewer than three groups. Teachers rated girls (M = 46.67, SD = 8.84) higher than boys (M = 41.96, SD = 8.14). Similarly, parents rated girls (M = 45.50, SD = 7.43) higher than boys (M = 42.09, SD = 10.05) (see Table 4.7) for Adaptive Skills scale.

Lastly, on the BSI scale across parent and teacher ratings, the Box's test for equality of covariances revealed that the assumption of homogeneity of covariance was not violated. Multivariate analyses indicated significant differences across genders on the multivariate combination of the parent and teacher ratings, Pillai's Trace = .09, $F(2, 138)$ = 6.707, p = .002, partial η^2 = .89. Univariate analyses revealed significant differences between the groups on [Teachers], $F(1, 139)$ = 7.726, p = .006, partial η^2 = .05, and [Parents], $F(1,139)$ = 10.74, p = .001, partial η^2 = .07. Post hoc tests were not performed for Gender because there are fewer than three groups. Teachers rated boys (M = 57.92, SD = 11.83) higher than girls (M = 52.10, SD = 11.88). Parents also rated boys (M = 58.51, SD = 12.11) higher than girls (M = 52.00, SD = 9.32) (see Table 4.7) for Behavior Symptom Index.

Table 4.7

MANOVA - Gender Differences for Parents' and Teachers' Ratings on the BASC-2

		Gender	N	M	SD	F(df)
Internalizing		1	50	53.38	12.58	2.845 (1, 140)
	Teacher	2	92	57.14	12.75	
	Parent	1	50	53.28	13.89	
		2	92	56.98	13.03	2.491 (1, 140)
Externalizing		1	50	49.44	8.97	
	Teacher	2	92	55.54	11.11	11.114 (1, 140)**
	Parent	1	50	50.26	7.59	
		2	92	56.13	11.89	9.958 (1, 140)*
Attention Problems		1	50	54.70	10.92	
	Teacher	2	92	59.17	9.98	6.084 (1, 140)*
	Parent	1	50	55.94	10.65	
		2	92	59.28	10.19	3.372 (1, 140)
Adaptive Skills		1	48	46.67	8.84	
	Teacher	2	91	41.96	8.14	9.907 (1, 137)*
	Parent	1	48	45.50	7.53	
		2	91	42.09	10.04	4.265 (1, 137)*
Behavior Symptom Index	Teacher	1	49	52.10	11.87	7.72 (1, 139)*
		2	92	57.92	11.82	
	Parent	1	49	52.00	9.32	
		2	92	58.51	12.11	10.74 (1, 139)**

Note. *p<.05, **p<.001

Summary Research Question 4

Overall the results of the analysis suggest that there were gender differences for all scales except Internalizing scale. On the Attention Problems scale there was no difference in parent ratings. The hypothesis for girls having higher ratings on the Internalizing scale was not supported. The other hypotheses that there would be gender differences on the other scales were supported. Boys were rated higher on all scales except on the Adaptive Skills scale.

Table 4.8 lists the hypotheses related to each research question and whether or not

it was supported. For Research Questions 1 and 2 where multiple imputed data sets were

used, qualitative comparisons were made across the original data set and the five data sets

to compare results. Similar results were found across data sets for the correlations that

were calculated for Research Question 1. The analyses for Research Question 2 included

a stepwise regression model and across the original data set and the multiple imputed data

sets the same predictors (informants) were found across models. However, the final step

in the model did not always include the same predictor as the ones found for the pooled

results. This suggests that qualitatively, there were comparable findings across the

original, imputed and pooled results.

Table 4.8

Summary of Hypotheses for Research Questions

Hypotheses	Supported or Not Supported
Level of Agreement	
(a) Mothers and Fathers will have higher levels of agreement on all scales except Internalizing.	Not Supported
(b) All teachers will have higher levels of agreement on all scales except on the Internalizing Scale.	Supported
(c) Across parents and teachers, raters will have higher levels of agreement on Externalizing, Behavioral Symptom Index and Attention Problems scales.	Supported
(d) Across parents and teachers, raters will have lower levels of agreement on Internalizing and Adaptive skills scales.	Supported
Predictive Ability	
(a) Parents' scores will more often predict the final diagnosis compared to teachers for internalizing diagnoses.	Supported
(b) Parents' scores will be sufficient to predict the diagnosis of an internalizing disorder.	Not Supported
(c) Teachers' scores will more often predict the final	Not Supported

diagnosis compared to parents' scores on externalizing and attention related diagnoses.	
(d) Teachers' scores will be sufficient to predict the diagnosis of an externalizing disorder and in the diagnosis of attention problems.	Not Supported
(e) Parents' and Teachers' scores will be needed to predict no diagnosis or "other" diagnoses (i.e. Learning Disorders, etc.).	Not Supported
Positive vs. Problem Behavior	
(a) Parents will report more positive behavior as measured by the Adaptive Skills scale than teachers.	Not Supported
(b) Teachers will report more problem behaviors as measured by the Behavioral Symptom Index.	Not Supported
Gender Interaction	
(a) There will be gender differences in informant ratings based on the gender on the Behavioral Symptom Index.	Supported
(b) Boys will have more problem behaviors on the Externalizing and Attention Problems scales compared to girls.	Supported
(c) Girls will have more problems on the Internalizing Scale compared to boys.	Supported

Other Variables

Other demographic information were collected from the evaluations (Age, Birth Order, Type of family, Teacher length of time known and Teacher subjects taught). The following is a brief summary of the analysis and results for these variables across ratings of parents and teachers for subscales in the original data set.

Age was separated into two groups; pre-adolescent (ages 6-12) and adolescent (ages 13-17) was compared across all subscales using MANOVA, no significant results were found. Similarly conducting ANOVA tests for length of time known for teachers (group1= < 6 months and group 2 = > 6 months) there were no significant differences in teacher's ratings across subscales. In examining subject of the teacher (academic vs. non-academic), no significant differences were found across the scales.

66

Type of family was examined by having two groups: intact families and other families (blended families, single parents, grandparents, adoptive parents and same sex parents). Using MANOVA tests, the multivariate and univariate tests were significant specifically for teachers, $F(1, 133) = 12.87$, $p = .000$, partial $\eta^2 = .09$ on the Externalizing scale. Comparing the means for Teachers' ratings for other families ($M = 58.62$, $SD = 11.68$) and intact families ($M = 51.51$, $SD = 9.34$) indicate that Teachers rated other families (blended, grandparents, and adoptive parents, single parents) higher than intact families on the Externalizing scale. On the Adaptive Skills multivariate tests were not significant; however, significant results were found for teachers' ratings on univariate tests, $F(1, 130) = 4.00$, $p = .047$, partial $\eta^2 = .03$. Teachers rated intact families ($M = 44.49$, $SD = 8.56$) higher on Adaptive Skills than other families ($M = 41.15$, $SD = 7.90$).

Birth order was examined as each child was coded into one of the following groups: oldest, middle, youngest, and only child (twins were not included in the data analysis). Step-brothers and sisters and half-siblings were included in determining birth order if they were living in the same household of the child. Using MANOVA tests, significant results were found for teachers' ratings on the BSI scale for only children, $F(3, 130) = 4.33$, $p = .006$, partial $\eta^2 = .09$. Teachers rated only children ($M = 61.54$, $SD = 10.96$) higher on BSI than oldest, middle or youngest children. Post hoc Scheffe tests were then computed to determine specifically which groups differed from which other groups. Post hoc Scheffe tests revealed differences between oldest and only children with only children having higher scores ($p = .020$).

CHAPTER 5

DISCUSSION

This chapter first presents an overview of the current study and a summary of the

variables in the study. Additionally, the applicability of these concepts in assessing socio-

emotional disorders is considered. In the subsequent section the findings are discussed

and incorporated with the current literature; followed by the limitations of the study.

Lastly, the implications and recommendations for future research are suggested.

Summary of Study

The purpose of this study was to examine the 1) the level of agreement between

multiple informants 2) the ability of each informant to predict the final diagnosis 3) the

difference in informants' ratings of positive behavior compared to problem behavior and

4) the gender difference in the informants' ratings on the BASC-2 to aid in assessing for

socio-emotional disorders. Largely, the objective of this study was to consider these

variables on the BASC-2 to ensure the appropriate use of behavior checklists when

diagnosing socio-emotional disorders in children.

Agreement or lack thereof between informants is relevant when behavior

checklists are used as part of a comprehensive psychological evaluation in assessing

children and adolescents (Merrell, 2008). It is more critical when it is used as a stand-

alone tool to diagnose certain disorders such as ADHD. Behavior checklists such as the

BASC-2 are attractive in managed care settings when there is a need for cost and time-

efficient assessment tools (Achenbach & Ruffle, 2000). However, these checklists need to be used appropriately and interpreted in light of differences in informants' ratings. Prior research has suggested that informants' rating vary by the construct of the scale is measuring (i.e. internalizing vs. externalizing); by types of informants (mother vs. fathers); and the setting of the informant (home vs. school) (Duhig, et al., 2000; Greitens et al. 2004). Studies have found that across settings (home vs. school) there are low levels of agreement and moderate to higher levels of agreement within same settings (home/school) (Christensen et.al, 1992). Higher levels of agreement have been found for informants when rating externalizing behavior and lower levels of agreement when rating internalizing behavior (Duhig, et.al, 2000; Grietens et al.2004). In two meta-analytical studies on the CBCL, the authors found higher levels of correlation (.60) for informants within the same setting (i.e. mothers and fathers). Much lower correlations were noted in the studies, when comparing teachers and parents across (home vs. school) .28 (Achenbach, McConaughy & Howell, 1987; Duhig, et al., 2000). On the BASC-2 a few studies have had similar results to the CBCL (Van Slyke, 2008; Neill, 2002). Much of the research has been focused on the CBCL even though the BASC-2 is just as widely used (Merrell, 2008; Rescorla, 2009). To increase the research base on the BASC-2, this study examined the level of informant agreement among four informants (mothers, fathers, and two teachers) across two settings (home and school).

Although constructs within behavior checklists are designed to aid in the diagnosis of disorders based on the DSM-IV-TR (2000), the predictive ability of checklists needed to be examined. Research studies have shown that raters are more

likely to predict externalizing disorders such as ADHD, ODD, and Conduct Disorders (Granero, Lourdes, Ezpeleta, Osa & Domenech, 2009). On the other hand, predictive ability for internalizing disorders has not been as evident (Dierker et al., 2001). Therefore this study further investigated the extent to which informants predict a diagnosis and whether it is necessary to have multiple informants.

A unique feature to the BASC-2 is the inclusion of positive behavior as measured by the Adaptive Skills subscale within the checklist (Rescorla, 2009). Research has primarily focused on problem behavior, but little is known about informants' ratings on positive behavior. A few studies have suggested that parents are seen as the most reliable source to report positive behavior (Loeber, Green, & Lahey, 1990; Phares, 1996; Hines & Paulson, 2006). As counseling psychology places the emphasis on strength-based perspectives, it is important to investigate the clinical utility of the Adaptive Skills scale. Hence, this study compared informants' ratings across positive behavior and problem behavior.

Lastly, gender effects on informant ratings have had varying results based on different studies. Achenbach et al. (1987) found that gender was not a factor in informants' ratings; however, Duhig et al. found gender differences for ratings of externalizing behaviors. Specifically, mothers and fathers rated more externalizing behaviors for boys than girls. On the BASC-2, agreement between raters differed based on gender for specific behaviors with more boys reported as having externalizing behavior and girls having internalizing behavior (Van Slyke, 2008; Gross-Tsur et al.,

70

2006; van der Ende & Verhulst, 2005; Sourander & Helstela, 2005). As studies have had mixed results regarding gender effects, this study sought to further examine gender differences across informant ratings.

Current Findings Related to Literature

A chart-review study was conducted on assessments of children and adolescents from a university based outpatient clinic. Out of 471 assessments, 100 were completed using the BASC-2 between the years of January 2000 to May 2011. The sample of evaluations included children and adolescents, age 6-17 who were referred and evaluated for a learning disorder or socio-emotional disorder. The evaluations for review included predominately White (92 percent) children and adolescents from intact families (71.3 percent). The following diagnoses were made in the evaluations Mood Disorders (Major Depressive Disorder, Dysthymia, and Depression Disorder NOS), Anxiety Disorders (Generalized Anxiety Disorder, Anxiety Disorder NOS), Attention Deficit/Hyperactivity Disorder (all types); Disruptive Behavior Disorders (Conduct Disorder and Oppositional Defiant Disorder); and other disorders (Learning Disorders, Reactive Attachment Disorder, and Aspergers Disorder).

The first research question addressed the level of agreement between ratings of different informants (mothers, fathers, and two teachers). Specifically, the levels of agreement between informants were examined across five subscales of the BASC-2: Internalizing, Externalizing, Adaptive Skills, Behavioral Symptom Index and Attention Problems. Consistent with previous studies on behavior checklists mothers and fathers had a strong relationship and moderate levels of agreement on all scales including

internalizing (Achenbach et al. 1987; Duhig et al., 2000). Although it was hypothesized in the present study that levels of agreement between mothers and fathers would be lower for internalizing compared to other scales, this was not supported. One study found that internalizing behavior was comparable to ratings on other scales (Youngstrom, Loeber, & Stouthamer-Loeber, 2000). Similar to other studies, teachers had fair to moderate levels of agreement on all scales except Externalizing which had higher levels of agreement (Van Slyke, 2008; Woo et al. 2007; Achenbach et. al, 1987). Lastly, there were higher levels of agreement between parents and teachers for Externalizing and Attention Problems scale and lower levels of agreement on Internalizing and Adaptive Skills which was supported in other studies (Fergusson & Horwood, 1993; Frauenglass & Routh, 1999; Hay et al., 1999; Hinshaw & Nigg, 1999; Kumpulainen et al., 1999; McConaughy, 1992; Merrell, 1999; Phares, 1997; Routh, 1990; Silverman & Rabian, 1999). As hypothesized and supported in this study, the level of agreement for parents and teachers for Behavioral Symptom Index (for total problems) were similar to results in other studies (Verhulst & Akkerhuis, 1989; Winsler & Wallace, 2002; Yeh & Weisz, 2001; Youngstrom et al., 2000).

The examination of the predictive ability for informants' ratings on the final diagnosis had mixed results. Studies have found that externalizing disorders and attention disorders are more easily predicted by informants' ratings (Tripp, Schaughency, & Clarke, 2006). Best practice for the use of behavior checklists is to include multiple informants in assessing children and adolescents (Grietens et al., 2004; Treutler & Epkins, 2001; Merrell, 2008). However, this can be time consuming and costly. The

results of the current study suggest that parents' ratings on the Internalizing and Attention Problems scale are sufficient in predicting the categories of internalizing and attention disorders which has been found in other studies (Phares, 1996; Stanger & Lewis, 1993). Even though teachers were included on overall models that were significant, they were not significant as individual predictors. This is inconsistent with other studies that found that teachers were predictive of attention-related disorders (Scholte, Van Berckelaer-Onnes & Van der Ploeg, 2008; Granero, Lourdes, Ezpeleta, Osa & Doménech, 2009).

Another aspect in this study was to examine whether informants' perceptions differed on positive behavior compared to problem behavior. To the author's knowledge this is the first study that compared these two scales across parents and teachers. Despite the hypothesis that there would be a difference between raters on these scales, the present study found that parents and teachers rated positive behavior and problem behavior similarly. There was however, a difference in agreement between raters on Adaptive Skills, with higher levels of agreement between ratings of mothers and fathers than the two teachers. One study (Acuna, 2003) that examined positive behavior (as measured by Adaptive Skills) found that parents rated positive behavior lower than problem behavior, but did not compare this between raters.

Research Question 4 addressed gender interaction effects on informants' ratings were also considered. The impact of the child's gender on informants' ratings has had mixed results depending on the study (Achenbach et. al, 1987; Duhig et.al, 2000). Numerous studies have found that informants rate children differently based on their gender (Van Slyke, 2008; Gross-Tsur et al., 2006;). The findings of the current study

73

indicate that parents and teachers rate boys higher than girls on the Externalizing scale, the Adaptive Skills scale and the Behavior Symptom Index scale. Additionally, teachers rated boys higher than girls on the Attention Problems scale. These results are consistent with findings in other studies (Carbonaro, 2009; Greitens et al., 2004). Surprisingly, no gender differences were found for the Internalizing scale which is inconsistent with other studies that found girls to have higher ratings than boys for informants' ratings on internalizing behavior (Treutler & Epkins, 2003; Lee, Elliot & Barbour, 1994; Sawyer, Baghurst, & Mathis, 1992). This may have been as a result of having a limited number of evaluations with an internalizing diagnosis.

Other variables not specifically hypothesized in this study were also examined. No significant findings were found for informants' ratings across subscales for age, length of time known and subject taught by teacher. Interestingly, differences in ratings were found for teachers for the type of family (intact vs. other) and birth order. Other families included blended families, single parents, grandparents, adoptive parents and same sex parents. Teachers rated "other" families higher than intact families on the Externalizing scale and lower on the Adaptive Skills scale. A few studies have found that parents' ratings across scales differed based on the type of family (Kurtz, 1995; Hilton, 1998); however, no studies were located that studies teachers' ratings across types of families. In studying the birth order the results of the current study indicated that teachers rated only children as having higher total problems compared to oldest, middle and youngest children. This may be reflective of whether teachers' perceive being an only

74

child or not being part of an intact family as a problem (McGee, Silva, & Williams, 1983).

Conclusions

In sum, the findings in this study examine factors that should be taken into consideration when using multiple informants in the assessment of children and adolescents. Specifically in examining informant agreement, predictive ability, differences in ratings across positive behavior and problem behavior and gender differences in informants' ratings, evaluators can make an informed decision in incorporating the BASC-2 in the assessment of children and adolescents. This study supports previous findings in the use of multiple informants across settings (home vs. school) as additional information that adds to the assessment data gathered during the evaluation of the child (Schroder, Hood & Hughes, 2010). Therefore, these factors examined in this study should be considered when interpreting ratings of multiple informants.

In accordance with other research on behavior checklists, this study on the BASC-2 yielded similar results for level of agreement between informants across settings (home vs. school) (Achenbach et.al, 1987; Duhig, et. al, 2000). Informants (mothers, fathers and teachers) from the same setting (home/school) have higher levels of agreement; however ratings were much lower between parents and teachers which reflect the differences in observed behavior across settings (home vs. school). These differences were noted in teachers' ratings of externalizing behavior which is readily observed by teachers when children "act out." More often teachers do not have significant amounts of one-on-one

contact with children limiting their knowledge of the child's behavior (Mclaughlin, 2004). As such, parents and teachers differed in ratings for Internalizing and Adaptive Skills subscales. These scales are geared towards knowing the internal world of the child and having more one-on-one interactions with the child as parents often do. In this study, fathers' ratings predicted internalizing and mothers' ratings predicted attentional disorders (compared to the absence of these disorders) lending support that parents have more knowledge about their children's behaviors than teachers. Although no differences were found between parent and teacher ratings of positive and problem behavior, there were differences in the level of agreement between parents and teachers on Adaptive Skills which suggest that further exploration is needed in this area. Gender differences were found among raters across the subscales. Teachers rated boys higher on Externalizing, Attention Problems, and Behavior Symptom Index. Parents and teachers rated girls higher on Adaptive Skills which may suggest that positive behavior manifests itself differently across genders or informants' may be biased in how they rate boys and girls on specific behaviors. This supports findings in previous studies that the gender of the child influences ratings of informants (Duhig, et al. 2000; Grietens et al., 2004).

Overall, this study extends the body of research on the BASC-2, to allow evaluators to take into consideration the above factors when interpreting ratings between informants. This is more critical when behavior checklists are used as stand-alone tools for the diagnosis of disorders. The current study explored factors that are integral to interpreting behavior checklists and more specifically the BASC-2. As psychological assessment of children and adolescent moves in the direction of using more cost effective

and time efficient tools – if not already, behavior checklists will soon lead the way

(Achenbach & Ruffle, 2000).

Limitations

Previous studies have found similar findings to those in the present study; however, there are several limitations to this study that should be taken into consideration. Most significantly, the data used in this study included a small sample size and contained a significant number of missing values. Additionally, a stepwise regression was used for analysis which is sample dependent. This limits the generalizability of the findings outside of the current sample. Although there were significant and non-significant findings, replication of this study is needed to confirm these findings.

Missing data was resolved by the use of multiple imputation methods which has been available for awhile, but is relatively new to the field of psychology. In this study, given that stepwise logistic regression, independent t-tests, and MANOVA were used for analysis; the use of multiple imputation strategy provided some unique challenges to data analysis. Thus some research questions used the original data set and others used the imputed data sets which could result in inconsistent findings.

Another limitation to this study is that the study focused on noting relationships and differences, but causality for the differences was not considered. For example, why do mothers' ratings predict internalizing and attention-related disorders? This study could have limited the focus by addressing the factors that cause the differences rather than just naming them. Lastly, this was a post-hoc study and therefore variables were not controlled prior to data collection. The sample included predominately White children

77

and adolescents in two-parent homes which limit the conclusions that can be drawn to other diverse populations.

Implications

The limitations of this study were presented as a backdrop prior to the discussion of implications to acknowledge that the inferences regarding the findings in this study are confined to this sample. Behavior checklists are used as a standard assessment tool in most evaluations; checklists are easy to administer and provide a wealth of information that can be used in diagnosing children and adolescents (Merrell, 2008). However, the use of these checklists in the evaluation process varies across assessors. Some use it as a small part of the evaluation and others rely solely on the information that the checklist provides. Managed care's emphasis on providing the best service with the least amount of resources has moved assessment into the direction of diagnosing children based on the most cost/time-efficient tools (Achenbach & Ruffle, 2000).

Given the findings on level of agreement among informants in this study, the use of multiple informants is needed in the assessment process (Schroeder, Hood, & Hughes, 2010). Each informant adds to the picture of what is going on with the child (Duhig et. al, 2000). If there are two informants from the same setting (i.e. mothers and fathers or multiple teachers), the ratings can be compared for agreement by evaluating the t-scores. If however, the informants are from different settings (home vs. school), noting the differences allows the evaluator to further assess for these differences (i.e. Is it the structure of the school? Is the teacher only noting externalizing behavior? etc.). The differences in level of agreement of ratings between parents and teachers on positive

behavior and internalizing behavior should shift evaluators' attention to these scales. For teacher ratings, externalizing scales can be compared depending on the referral question (i.e. referral for Disruptive Disorders). The differences in agreement between raters may be accounted for by variability in behavior at home and at school, errors related to the instrument and informant variables (i.e. bias towards the child, knowledge regarding child development, mental health of the informant etc.) (Fergusson& Horwood, 1993; Grietens et. al, 2004). Given these differences, a detailed analysis of the informants' profiles is needed – beyond examining t-score and clinical significance.

In the examination of predictive ability, parents (mothers and fathers) were more likely to appear in a model that was predictive of an internalizing diagnosis. For attention-related diagnosis mothers and teachers were more likely to appear in a model. These findings suggest that when multiple informants are not available it is best to obtain ratings from mothers and one teacher (Grietens et. al, 2004) when assessing for attention-related diagnosis. More informants provide additional information that may be lacking if a brief evaluation is conducted. However, taking into consideration the cost and time constraints of evaluators, it seems that parents may be the best informants. In particular, it is important to note the differences in the level of agreement between informants and across settings (home vs. school). As evident in this study, mothers and fathers tend to rate children similarly, if there is a big discrepancy in ratings, possible reasons for these differences should be explored rather than taking the ratings at face value.

There were no significant findings for the "no diagnosis" and "other" diagnostic categories in this study. The "other" category included a hodgepodge of diagnoses with

79

the majority falling into the learning disorders category. As such the BSI scale would not reflect these diagnoses and therefore not predict this category. However, the Adaptive Skills scale would have reflected lower scores for those that did not have a diagnosis and be predictive of no diagnosis, but this was not found. This is consistent with the lack of significant findings on informant ratings between problem behavior and positive behavior. Further examination of these variables could help shift the focus of the BASC-2 from a tool used for assessing problem behavior, to a tool also used in assessing positive behavior. The lack of significant findings may be accounted for by the type of analysis conducted or the coding of the variables.

As gender differences were found, evaluators need to take into consideration the gender of the child and be aware that boys may have inflated scores for identifying externalizing, attention problems and total problems. An interesting finding is that girls were rated higher on the Adaptive Skills scale which may suggest that informants view boys as having more problems and girls as having more positive behavior. Further behaviors may be expressed differently between boys and girls and this may account for the gender difference (Mclaughlin, 2004).

<div align="center">**Future Directions**</div>

As in any research study, the results and findings lead to more questions than answers. Overall, future research on all these factors needs to be replicated with a larger sample size and include a more diverse population as these were significant limitations in this study. Future studies should consider the following as it relates to the variables examined in this study:

1. In examining level of agreement between informants, the next step would be to tease out the reasons for the differences. Studies have been conducted on depressive mothers (McFarland & Sanders, 2003); however further research needs to consider teacher variables and biases for discrepancies in ratings

2. Given the small sample size and the type of analysis (stepwise regression) used, this study should be replicated with a larger sample size. Additionally, it would be helpful to understand why fathers are better predictors of an internalizing diagnosis and why mothers are predictive of attention-related diagnosis which may be better investigated using a qualitative study. If an "other" category is used, the School Problems scale on the teacher report should be gathered to determine predictive ability of learning disorders. Additionally as there was limited data on the externalizing category, further study of the informants' ability to predict an externalizing diagnosis should be investigated.

3. Positive behavior and problem behavior need to be further examined (with different analyses or coding of variables) to fully understand how informants' rate these two scales. This would be beneficial especially for counseling psychologists who are rooted in strength based perspectives.

4. The next step in examining gender differences would be to understand why informants rate boys and girls differently on certain scales? This would help in understanding gender biases that exist in informant ratings.

5. Other variables in this study (age, birth order, family type, length of time known by teacher, and subject taught by teacher) should be further explored as possible factors that may influence informants' ratings in particular birth order or type of family (intact vs. other families) as significant findings were found in this study.

6. As this study used multiple imputation as a method to address missing data, a comparison of these results to complete data sets would provide further credibility for use of multiple imputation for statistical analyses within psychological research.

REFERENCES

Achenbach, T.M. (2010). *ASEBA bibliography on the web.* Retrieved from

http://www.asebabib.org/

Achenbach T.M., Becker, A., Dopfner, M., Heiervang, E., Roessner, V., Steinhausen, H.C., &

Rothenberger, A. (2008). Multicultural assessment of child and adolescent

psychopathology with ASEBA and SDQ instruments: Research findings, applications,

and future directions. *Journal of Child Psychology and Psychiatry* 49(3), 251–275.

Achenbach, T.M., Dumenci, L., & Rescorla, L.A. (2002). Ten-year comparisons of problems and

competencies for national samples of youth self, parent, and teacher reports

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.

Achenbach, T.M. & Howell, C.T. (1993). Are American children's problems getting worse? A

13-year comparison. *Journal of American Academy of Child and Adolescent Psychiatry,*

32 (6), 1145-1154.

Achenbach, T.M., & Rescorla, L.A. (2001). Manual for the ASEBA School-Age Forms &

Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth,

and Families.

Achenbach T.M. & Ruffle T.M. (2000). The Child Behavior Checklist and related forms for

assessing behavioral/emotional problems and competencies. *Pediatrics In Review /*

American Academy Of Pediatrics, 21(8), 265-271.

Acuna, T.L.(2003). Patterns and stability in clinical and adaptive ratings of emotionally disabled children. (Doctoral Dissertation). Available from *Dissertation Abstracts International Section A: Humanities and Social Sciences, 64*(6-A), 2039.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders* (1st ed.). Washington, DC: Author.

American Psychological Association (2003b). *Report of the APA Task Force on Psychology's Agenda for Child and Adolescent Mental Health.* Washington, DC: Author. Retrieved from http://www.apa.org/pi/families/resources/child-mental-health.pdf

American Psychological Association. (2003a). Multicultural guidelines: Education, research, and practice. *American Psychologist, 58,* 377–402.

Angold, A., & Costello E.J. (1993). Depressive comorbidity in children and adolescents: empirical, theoretical, and methodological issues. *American Journal of Psychiatry, 150*(12), 1779-1791.

Aschenbrand, S. G., Angelosante, A.G., & Kendall, P. C., (2005). Discriminant validity and clinical utility of the CBCL With anxiety**-d**isordered youth. *Journal of Clinical Child & Adolescent Psychology, 34*(4), 735-746.

Barkley, R. A. (2006). *Attention Deficit Hyperactivitv Disorder: A handbook for diagnosis and treatment* (3rd ed.) New York: Guilford Publications.

Betz N. (2008). Advances in vocational theories. In Brown S. & Lent, R. (Eds.) *Handbook of Counseling Psychology* (4th ed.) (pp. 357-374). Hobokem, NJ: John Wiley & Sons.

Birmaher, B., Ryan, N.D., Williamson, D.E., & Brent D.A., Kaufman, J., & Dahl, R.E., Perel, J.,

  Nelson, B. (1996). Childhood and adolescent depression: a review of the past 10 years.

  Part I. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(11),

  1427-1439.

Bisson, J.I. & Sakhuja, D. (2006). Adjustment disorders. *Journal of Psychiatry, 5*(7), 240-242.

Brauner, C.B., & Stephens, B.C. (2006). Estimating the Prevalence of Early Childhood Serious

  Emotional/Behavioral Disorder: Challenges and Recommendations. *Public Health*

  Reports 121, 303-310.

Brown, S.D. & Rector C.C. (2008). Conceptualizing and diagnosing problems in vocational

  decision making. In Brown S. & Lent, R. (Eds.) *Handbook of Counseling Psychology* (4th

  ed.) (pp. 392-407). Hobokem, NJ: John Wiley & Sons.

Brook, W.S. (1996). *Congruence between child self-reports and those of their parents and*

  teachers on anxiety and depression: A preliminary investigation of parameters of

  agreement. (Doctoral Dissertation). Available from *Dissertation Abstracts International:*

  Section B: The Sciences and Engineering, 57(2-A), 0575. Retrieved from PsychINFO

  database. (1996-95016-073)

Buehler, C., Anthony, A., Krishnakumar, A., Stone, G., Gerard, J., & Pemberton, S. (1997).

  Interparental conflict and youth problem behaviors: A meta-analysis. *Journal of Child*

  and Family Studies, 6(2), 233-247.

Bulmer, M.G. (1979). Principles of statistics. Dover Publications.

Carbonaro, D. T. (2009). Cross-informant discrepancy analysis in children with internalizing and

      externalizing behavior disorders. *Dissertation Abstracts International: Section B: The*

      Sciences and Engineering,69(12-B), 7804.

Carr, A. (2008**).** Depression in young people**:** description**,** assessment and evidence**-**based

      treatment**.** *Developmental Neurorehabilitation, 11*(1), 3-15.

Center for Mental Health in Schools at UCLA (2008). Youngsters' mental health and

      psychosocial problems: What are the data? Los Angeles, CA: Author.

Chistensen, A., Margolin, M., & Sullaway, M. (1992). Interparental agreement on child behavior

      problems. *Psychological Assessment, 4*(4). 419-425.

Collet, B.R., Ohan, J.L., & Myers, K.M., Ten-year review of rating scales V: Assessing

      attention-deficit/hyperactivity disorder. *Journal of American Academy of Child and*

      Adolescent Psychiatry,42(9), 1015-1037.

Conner, D., Edwards, G., Fletcher, K., Baird, J., Barkey, R., Steingard, R. (2003). Correlates of

      comorbid psychopathology in children with ADHD. *Journal or Child & Adolescent*

      *Psychiatry, (42)*2, 193-200.

Cooper, C. & Gottlieb, M. (2000). Ethical issues with managed care: Challenges facing

      counseling psychology. *The Counseling Psychologist, (28)* 179-236.

Dababnah S. & Cooper J. (2006). Unclaimed children revisited working paper No.1: Challenges

      and opportunities in children's mental health a view from families and youth. *National*

      Center for Children in Poverty, 1-4. Retrieved from

      http://www.nccp.org/publications/pdf/download_9.pdf

Dierker, L.C., Albano, A.M., Clarke, G.N., Heimberg, R.G., Kendall, P.C.,& Merikangas, K.R., (2001). Screening for anxiety and depression in early adolescence. *Journal of American Academy of Child Adolescent Psychiatry*, *40*, 929–936.

Doyle, A., Ostrander, R., Skare, S., Crosby, R. D., & August, G. (1997). Convergent and criterion-related validity of the Behavior Assessment System for Children—Parent Rating Scales. *Journal of Clinical Child Psychology, 26,* 276-284.

Duhig, A.M., Renk, K., Epstein, M.K., & Phares, V. (2000). Interparental agreement on Internalizing, Externalizing, and Total Problems: A meta-analysis. *Clinical Psychology: Science and Practice*, *7,* 435–453.

Farone, S. V., Althoff, R. R., Hudziak, J. J., Monteaux, M., Bierderman, J. (2005). The CBCL predicts DSM bipolar disorder in children: a receiver operating characteristic curve analysis. *Bipolar Disorders, 7*(6), 518-524.

Ferdinand, R. F. (2007). Validity of the CBCL/YSR DSM-IV scales anxiety problems and affective problems. *Journal of Anxiety Disorders, 22*(1), 126-134.

Fergusson, D. M., & Horwood, L. J. (1993). The structure, stability and correlations of the trait components of conduct disorder, attention deficit and anxiety/withdrawal reports. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 34*, 749-766.

Fischer, M., Barkley, R. A., Smallish, L., & Fletcher, K. (2004). Hyperactive children as young adults: Deficits in inhibition, attention, and response perseveration and their relationship to severity of childhood and current ADHD and conduct disorder. *Developmental Neuropsychology*, 27, 107-133.

Frick, P.J., Barry, C.T., & Kamphaus, R.W. (2009). *Clinical assessment of child and adolescent personality and behavior.* New York: Springer.

Fox, J. (1997) *Applied regression analysis, linear models, and related methods.* Thousand Oaks, CA: Sage Publications.

Granero, Ezpeleta, Osa, & Domenech (2009). Predictive validity of symptoms-based approaches to the adjustment of high-risk children at school and in the community. *European Journal of Psychological Assessment,(25)*4, 274–284.

Ginzburg, S. G. A. (2009). *Cross-informant agreement between parents and adolescents in 20 cultures.* (Doctoral Dissertation) *Available from Dissertation Abstracts International: Section B: The Sciences and Engineering, 70*(4-B), Retrieved from PsychINFO database. (2009-99200-408)

Greenbaum, P. E., Dedrick, R. F., Prange, M. E., & Friedman, R. M. (1994). Parent, teacher, and child ratings of problem behaviors of youngsters with serious emotional disturbances. *Psychological Assessment*, *6,* 141–148.

Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., Van Assche, V., Ghesquiere, P., Hellinckx, W. (2004). Comparison of mothers', fathers', and teachers' reports on problem behavior in 5- to 6-year-old children. *Journal of Psychopathology and Behavioral Assessment, 26*(2), 137-146.

Gross-Tsur, V., Zedek, S., & Goldzweig, G., Landau, Y.E., Berger, I., Shalev, R.S. (2006). The impact of sex and subtypes on cognitive and psychosocial aspects of ADHD. *Developmental Medicine and Child Neurology, 48,* 901-905.

Hay, D. F., Pawlby, S., Sharp, D., Schm¨ucker, G., Mills, A., Allen, H., & Kumar, R. (1999). Parents' judgements about young children's problems: Why mothers and fathers might disagree yet still predict later outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40,* 1249–1258.

Hines, A.R. & Paulson, S.E. (2006). Parents' and teachers' perceptions of adolescent storm and stress: relations with parenting and teaching styles. *Adolescence, 41*(4), 597-614.

Hilton, J. (1998). Comparison of parenting and children's behavior in single-mother, single-father, and intact families. *Journal of Divorce & Remarriage,29,* (3/4), 23-54.

Hodgman C.H. (2007). Adjustment disorder. In Garfunkel, L.C., Kaczorowski, J., & Christy, C. *Mosby's pediatric clinical advisor: Instant diagnosis and treatment.* St. Louis, Missiouri: Mosby.

Hughes, C.W. & Melson, A.G. (2008). Child and adolescent measures for diagnosis and screening. In Rush, A. J., First, M. B., & Blacker, D. (Eds.) *Handbook of psychiatric measure*s (3rd ed.) (pp.251-308). Arlington, VA: American Psychiatric Publishing, Inc.

Jarratt, K., Riccio, C. and Siekierski, B. (2005). Assessment of attention hyperactivity disorder (ADHD) using the BASC and BRIEF. *Applied Neuropsychology,12*(2), 83-89.

Kamphaus, R.W., Petoskey, M.D., Cody, A.H., Rowe, E.W., Huberty, C.J. and Kamphaus, C.R. (1999). A typology of parent rated child behaviour for a national US sample. *Journal of Child Psychology and Psychiatry,40* 607–616.

Keiley, M.K., Lofthouse, N., Bates, J.E., Dodge, K.A., & Pettit, G.S. (2003) Differential risks of co-varying and pure components in mother and teacher reports of externalising and

internalising behavior across ages 5 to 14. *Journal of Abnormal Child Psychology 31,*
267-283.

Kessler, R.C., Avenevoli, S., & Merikangas, K.R. (2001). Mood disorders in children and
adolescents: An epidemiologic perspective. *Biological Psychiatry,49*(12), 1002-1014.

Kimel, L.K., (2009). Phenomenology of anxiety and fears in clinically anxious children with
autism spectrum disorders. *Dissertation Abstracts International: Section B: The Sciences
and Engineerin, 69*(8-B), 5033.

Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic
children: Correspondence among child, parent, and teacher reports. *Journal of Child
Psychology and Psychiatry,34,* 991–1006.

Konold, T.R. & Pianta, R.C. (2007). The influence of informants on ratings of childrens'
behavioral functioning: A latent variable approach. *Journal of Psychoeducational
Assessment,25,*(3), 222-236.

Kurtz, L. (1995). The relationship between parental coping dtrategies and children's adaptive
processes in divorced and intact families. *Journal of Divorce & Remarriage, 24,* (3/4),
89-110.

Larson,K., Russ, S. A., Kahn, R. S., & Halfon, N. (2007). Patterns of comorbidity, functioning
and service use for US children with ADHD. *Pediatrics,127*(3), 462-470.

Lee, S. W., Elliott, J., & Barbour, J. D. (1994). A comparison of cross informant behavior ratings
in school-based diagnosis. *Behavior Disorders*, *19,* 87–97.

Lett, N. J., & Kamphaus, R. W. (1997). Differential validity of the BASC Student Observation
System and the BASC Teacher Rating Scale. *Canadian Journal of School Psychology*,

13, 1-14.

Lichtenberg, J. W., Goodyear, R.K., & Genther, D.Y. (2008). The changing landscape of professional practice in counseling psychology. In Brown S. & Lent, R. (Eds.) *Handbook of Counseling Psychology* (4th ed.) (pp. 21-37). Hobokem, NJ: John Wiley & Sons.

Little, R.A. (1989). Testing the Equality of Two Independent Binomial Proportions. *The American Statistician, 43(*4), 283–288.

Loeber, R., Green, S.M., Lahey, B.B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *Journal of Clinical Child & Adolescent Psychology, 19*(2), 136-143.

Lopez, S.J. & Edwards, L. M. (2008). The interface of counseling psychology and positive psychology: Assessing and promoting strengths. In Brown S. & Lent, R. (Eds.) *Handbook of Counseling Psychology* (4th ed.) (pp. 86-102). Hobokem, NJ: John Wiley & Sons.

Loughran, S. (2003). Agreement and stability of teacher rating scales for assessing ADHD in preschoolers. *Early Childhood Education Journal, 30*(4), 247-253.

Martens, H.M. (2006). Factors contributing to parent/child discrepancies on the behavior assessment system for children. (Doctoral Dissertation). Available from *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67*(2-B), 1157. Retrieved from PsychINFO database. (2006-99016-151)

Mclaughlin, V. (2004). Comparison of teacher ratings and direct observations of children with problem behaviors. (Doctoral Dissertation). *Dissertation Abstracts International Section A: Humanities and Social Sciences,64*(10-A), 3594

McGee, R., Silva, P., & Williams, S. (1983). Parents' and teachers' perceptions of behaviour

    problems in seven year old children.*International Journal of Disability, Development and*

    Education, 30(2), 151-161.

Mental Health Commission (2003). *A final report of the President's New Freedom Commission*

    on Mental Health. Washington, D.C.: Mental Health Commission. Retrieved from

    http://www.mentalhealthcommission.gov/index.html

Merrell, K.W. (2008). *Behavioral, social, and emotional assessment of children and adolescents*

    (3rd ed.). New York: Lawrence Erlbaum Associates.

Merrell, K.W., Streeter, A.L., & Boelter, E.W. (2001). Validity of the home and community

    social behavior scales: Comparisons with five behavior-rating scales. *Psychology in*

    Schools, 38,(4), 313-325.

Morgan, A.B.W. (1998). Childhood social withdrawal: The role of social skills in the relation of

    withdrawal and anxiety. (Doctoral Dissertation). Available from *Dissertation Abstracts*

    International: Section B: The Sciences and Engineering, 59(6-B), 3067. Retrieved from

    PsychINFO database. (1998-95024-264)

Najman, J. M., Heron, M.A., Hayatbakhsh, M.R., Dingle, K., Jamrozik, K., Bor, W.,

    O'Callaghan, M.J., Williams, G. M. (2008). Screening in early childhood for risk of later

    mental health problems: A longitudinal study. *Journal of Psychiatric Research, 42*(8),

    694-700.

Neill, D.L. (2002). Emotionally disturbed (ED) profiles in children and adolescents from parent,

    teacher, and self-reports on the BASC. (Doctoral Dissertation). Available from

   Dissertation Abstracts International: Section B: The Sciences and Engineering,62(12-B),

   5975.

O'Connell M.E., Boat T., Warner K.E.,(Eds.) (2009). *Preventing mental, emotional, and*

   behavioral disorders among young people. Washington: National Academies Press.

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance.

   Psychological Bulletin, 83, 579-586.

Papageorgiou, V., Efrosini, K., Dafoulis, V., & Vostanis, P. (2008). Differences in parents' and

   teachers' ratings of ADHD symptoms and other mental health problems. *European*

   Journal of Psychiatry,22(4), 200-210.

Pelkonen, M., Marttunen, M., Henriksson, M., & Lönnqvist, J. (2005). Suicidality in adjustment

   disorder--clinical characteristics of adolescent outpatients. *European Child and*

   Adolescent Psychiatry, 45(3), 174-80.

Phares, V. (1996). Accuracy of informants: do parents that mother knows best? *Abnormal Child*

   Psychology, 25(2), 165-171.

Pillai-Bartlett Morrison, D.F. (1976) Multivariate statistical methods. Second edition. New York:

   McGraw-Hill.

Querido, J. G., Eyberg, S. M., & Boggs, S. R. (2001). Revisiting the accuracy hypothesis in

   families of young children with conduct problems. *Journal of Clinical Child Psychology*,

   30, 253–261.

Rescorla, L. A. (2009). Rating scale systems for assessing psychopathology: The Achenbach

   System of Empirically Based Assessment (ASEBA) and the Behavior Assessment

   System for Children-2 (BASC-2). In Matson, J. L., Andrasik, F., & Matson, M. L. (Eds.),

Assessing childhood psychopathology and developmental disabilities (pp.117-149). New York, NY : Springer.

Reynolds, C. R. & Kamphaus, R. W. (2002). *The clinician's guide to the Behavior Assessment System for Children (BASC).* New York: Guilford.

Reynolds, C. R. & Kamphaus, R. W. (2004). *BASC-2 Behavior Assessment System for Children, second edition manual.* Circle Pines, MN: American Guidance Service

Reynolds, C.R. & Kamphaus, R.W. (1992). *Behavior Assessment System for Children: Manual.* Circle Pines, MN: American Guidance Service, Inc.

Rosenbaum, J.F., Hirshfeld-Becker, D.R., Henin, A., Hubley, S., LaCasse, S., Farone, S.V., & Bierderman, J. (2008). The child behavior checklist broad-band scales predict subsequent psychopathology: A 5-year follow-up. *Journal of Anxiety Disorders,22*(3), 532-539.

Runeson, B.S, Beskow, J., & Waern, M. (1996). The suicidal process in suicides among young people. *Acta Psychiatrica Scandinavica, 93,*(1), 35-42.

Rubin, Donald B. (1987) Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons.

Saklofske, D.S., Andrews, J.W., Janzen, H.L., Phye, G.D. (Eds.) (2002). *Handbook of psychoeducational assessment: A practical handbook.* New York: Academic Press.

Salbach-Andrae, H., Klaus, L., & Lehmkuhl, U (2009). Patterns of agreement among parent, teacher and youth ratings in a referred sample. *European Psychiatry, 24*, 345-351.

Salvia, J., Ysseldyke, J., & Bolt, S. (2009). *Assessment: In special and inclusive education.* Belmont, CA: Wadsworth.

Sawyer, M. G., Baghurst, P., & Mathias, J. (1992). Differences between informants' reports describing emotional and behavioral problems in community and clinic-referred children: A research note. *Journal of Child Psychology and Psychiatry*, *33,* 441–449.

Sayette , M., Mayne, T., & Norcross, J. (2010). Insider's guide to graduate programs in clinical and counseling psychology. New York: The Guilford Press.

Schafer, J.L. and Graham, J.W. (2002). Missing data: Our view of state of the art. *Psychological Methods* 7(2), 147-177.

Schafer, J. L. (1997). Analysis of incomplete multivariate data. Chapman and Hall.

Schlomer, G., Bauman, Sheri., & Card, N. A. Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57*(1), 1-10.

Scholte, E.M., Van Berckelaer-Onnes, I., & Van der Ploeg, J. D. (2008). A rating scale to screen symptoms of psychiatric disorders in children. *European Journal of Special Needs Education 23*(1), 47-62.

Schroeder, J., Hood, M., & Hughes, H. (2010). Inter-parent agreement on the syndrome   scales of the Child Behavior Checklist (CBCL): Correspondence and discrepancies. *Journal of Child & Family Studies,19*(5), 646-653.

Shulz, K.B. (2002). ADHD symptomology. The contribution of depression and anxiety measure to an accurate diagnosis. (Doctoral Dissertation). Available from *Dissertation Abstracts I International: Section B: The Sciences and Engineering, 62*(8-B), 3815. Retrieved From PsychINFO database. (2002-95004-172)

Silverman, W.K. & Ollendick, T.H. (2005). Evidence-based assessment of anxiety and its

    disorders in children and adolescents. *Journal Of Clinical Child And Adolescent*

    Psychology, 34(3), 380-411.

Smith, B., Barkley, R. A., & Shapiro, C. (2007). Attention deficit hyperactivity disorder. In E. J.

    Mash & R. A. Barkley (Eds.), Assessment of Childhood Disorders (4th ed.). New York:

    Guilford.

Smith, D. (2007). *Introduction to special education: Making a difference*. Needham Heights,

    Massachusetts: Allyn & Bacon.

Soper, D.S. (2011). The Free statistics calculators website, *Online Software*,

    http://www.danielsoper.com/statcalc/

Sourander, A. & Helstelä, L. (2005). Childhood predictors of externalizing and internalizing

    problems in adolescence: A prospective follow-up study from age 8 to16. *European*

    Child & Adolescent Psychiatry,14(8), 415-423.

Speight S.L., & Vera E.M. (2008). Social justice and counseling psychology: A challenge to the

    profession. In Brown S. & Lent, R. (Eds.) *Handbook of Counseling Psychology* (4th ed.)

    (pp. 54-67). Hobokem, NJ: John Wiley & Sons.

Spencer, T.J. (2006). ADHD and comorbidity in childhood. *Journal of Clinical Psychiatry,*

    *(67)*Suppl 8, 27-31.

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers and children on

    internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology*,

    22, 107–115.

Strain, J., & Diefenbacher, A. (2008). The adjustment disorders: The conundrums of the

    diagnoses. *Comprehensive Psychiatry,49*(2), 121-130.

Strain, J.J., Smith, G.C., Hammer, J.S., McKenzie, D.P., Blumenfield, M., Muskin, P.,

    Newstadt, G., Wallack, J., Wilner, A., & Schleifer, S. (1998). Adjustment disorder: a

    multisite study of its utilization and interventions in the consultation-liaison psychiatry

    setting. *General Hospital Psychiatry, 20*(3), 139-149.

Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father

    reports on children's behavior related to parents' psychological symptoms *and* aspects of

    parent–child relationships? *Journal of Abnormal Child Psychology*, *31,* 13–27.

Tripp, G., Schaughency, E., & Clarke, B. (2006). Parent and teacher rating scales in the

    evaluation of Attention-deficit Hyperactivity Disorder: Contribution to diagnosis and

    differential diagnosis in clinically referred children. *Developmental and Behavioral*

    *Pediatrics,(27)*3, 209-218.

U.S. Department of Health and Human Services (1999). *Mental health: A report of the Surgeon*

    General. Rockville, MD: U.S. Department of Health and Human Services, Substance

    Abuse and Mental Health Services Administration, Center for Mental Health Services,

    National Institutes of Health, National Institute of Mental Health. Retrieved from

    http://www.surgeongeneral.gov/library/mentalhealth/home.html

U.S. Department of Health and Human Services (2004). *National Health and Nutritional*

    Examination Survey. Ann Arbor, MI: National Center for Health Statistics. Retrieved

    from http://www.nimh.nih.gov/health/topics/child-and-adolescent-mental-

    health/index.shtml

U.S. Department of Health and Human Services (2007). *National Health Interview Survey.*
Washington, D.C.: National Center for Health Statistics. Retrieved from
http://www.cdc.gov/nchs/data/series/sr_10/sr10_239.pdf

U.S. Department of Health and Human Services (2008). *National Health Interview Survey.*
Washington, D.C.: National Center for Health Statistics. Retrieved from
http://www.cdc.gov/nchs/data/series/sr_10/sr10_244.pdf

Van Der Ende J., & Verhulst, F.C. (2005). Informant, gender and age differences in ratings of
adolescent problem behavior. *European Child and Adolescent Psychiatry, Supplement,
14*(3), 117-126.

Van Slyke, K.B. (2008). Assessing childhood difficulties: Comparing the SDQ and the BASC-2.
(Doctoral Dissertation). Available from *Dissertation Abstracts International: Section B:
The Sciences and Engineering, 68*(11-B), 7289.

Vaughn, M. L., Riccio, C. A., Hynd, G.W., & Hall, J. (1997). Diagnosing ADHD
(predominantly inattentive and combine type subtypes): Discriminant validity of the
behavior assessment system for children and the Achenbach parent and teacher rating
scales. *Journal of Clinical Child & Adolescent Psychology, 26*(4), 349-357.

Verhulst, F. C.,&Akkerhuis, G.W. (1989). Agreement between parents' and teachers' ratings of
behavioral/emotional problems of children aged 4–12. *Journal of Child Psychology and
Psychiatry*, *30,* 123–136.

Verhulst, F.C., Dekker, M.C., & van der Ende, J. (1997). Parent, teacher and self-reports as
predictors of signs of disturbance in adolescents: whose information carries the most
weight? *Acta Psychiatrica Scandinavica*, *96,* 75-81.

Volker, M. A., Lopata, C., Smerbeck, A. M., Knoll, V. A., Thomeer, M. L., Toomey, J.A., & Rodgers, J.D. (2010). BASC-2 PRS Profiles for Students with High-Functioning Autism Spectrum Disorders. *Journal of Autism & Developmental Disorders,* 40(2), 188-199.

Waggoner, C.E. (2005). Comparison of the BASC-2 PRS to the BASC PRS in a population of children and adolescents classified as HFA, Asperger disorder or PDD NOS including convergent validity. (Doctoral Dissertation). Available from *Dissertation Abstracts I International: Section B: The Sciences and Engineering, 66*(11-B), 6316. Retrieved From PsychINFO database. (2006-99010-163)

Watts, R. J. (2004). Integrating social justice and psychology. *The Counseling Psychologist, (32)*6, 855-865.

Woo, B.S.C, Ng, T.P., Fung, D.S.S., Chan, Y.H., Lee, Y.P., Koh, J.B.K., & Cai, Y. (2007). Emotional and behavioural problems in singaporean children based on parent, teacher and child reports. *Singapore Med Journal, 48*(12), 1100-1112.

Wood, A. M., White, I. R., & Royston, P. (2008). How should variable selection be performed with multiply imputed data? *Statistics in Medicine, 27*(17), 3227-3246.

Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology*, *68,* 1038–1050.

Youngstrom, E., Meyers, O., Demeter, C., Youngstrom, J., Morello, L., Piiparinen, R., Feeny, N., Calabrese, J. R., & Findling , R. L.(2005). Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. *Bipolar Disorders, 7*(6), 507-517.

Appendix A

CHART REVIEW FORM

Cl # _____

Demographic information

Child Demographics:
DOB: _____ Age: _____
Gender: ☐ Female ☐ Male
Race: _____
Grade: _____

Household Composition:
Type of family: ☐ Intact family ☐ Blended family ☐ Other
Number of siblings: _____ Birth order: _____

Teacher Demographics

Informant	Subject Taught
Teacher 1	☐ Academic _____ ☐ Non-academic _____
Teacher 2	☐ Academic _____ ☐ Non-academic _____
Teacher 3	☐ Academic _____ ☐ Non-academic _____

T-Scores for the BASC Scales

Informants	Adaptive Skills	Internalizing	Externalizing	Attention Problems	Behavioral Symptom Index
Mother					
Father					
Teacher 1					
Teacher 2					
Teacher 3					

*At-Risk ** Clinically Significant ^Low Adaptability ^^ High Adaptability

<div align="center">Appendix B – Description of BASC-2 Subscales^</div>

Adaptive Skills	
Activities of Daily Living	The ability to act in a safe manner, perform simple daily tasks, and organize tasks, ex: "Attends to issues of personal safety"
Functional Communication	The ability to express ideas and communicate in ways that others can easily understand, ex: "Communicates clearly"
Leadership	Having skills to accomplish goals, ability to relate to others
Social Skills	Having skills associated with proper conduct when interacting with others, and exhibiting behaviors that encourage others, offer assistance, and display a concern for the well-being of others, ex: "Encourages others to do their best"
Externalizing	
Aggression*	The tendency to do physical or emotional harm to others or their property, ex:"Losses temper too easily"
Hyperactivity*	The tendency to act without thinking and be unable to wait one's turn in group activities, ex: "Interrupts others when they are speaking"
Conduct Problems	The tendency to behave in a socially deviant manner and exhibit disruptive behaviors, "Breaks the rules"
Internalizing	
Anxiety*	The tendency to excessively worry, develop unreasonable fears, self-deprecate, or be excessively nervous, ex: "Worries about making mistakes"
Depression*	The tendency to develop overly negative cognitions with regard to the self, the world, and the future, ex:"Seems lonely"
Somatization	The tendency to be overly sensitive and complain about relatively minor physical problems or ailments, and to over report the occurrence of various physical complaints, ex:"Gets sick"
Behavioral Symptom Index*	
Atypicality	The tendency to behave in ways that are considered odd or strange, ex: "Seems out of touch with reality"
Withdrawal	The tendency to avoid social contact and to lack interest in making contact in social settings, ex: "Makes friends easily"
Attention Problems	The tendency to be unable to maintain attention and become easily distracted from tasks requiring attention, ex: "Is easily distracted"

*Also includes these subscales for the Behavioral Symptom Index

Item examples from the Parent Report of the BASC-2

^ Adapted from Reynolds & Kamphaus (2004)

Appendix C - BASC-2 Scale Structure and Scoring System

Parent Rating Form	Teacher Rating Form	Scoring
Adaptive Skills Activities of Daily Living Functional Communication Leadership Social Skills	*Adaptive Skills* Study Skills Functional Communication Leadership Social Skills	T-score of 70 or above: Very High adaptive skills T-scores of 40-70: Average T-scores of 30-40: At-Risk T-scores of 30 or below: Clinically Significant
Externalizing Aggression Hyperactivity Conduct Problems	*Externalizing* Aggression Hyperactivity Conduct Problems	T-scores of 70 or above: Clinically Significant T-scores of 60-70: At-Risk T-scores of 40-60: Average T-scores of 40 or below: Low
Internalizing Anxiety Depression Somatization	*Internalizing* Anxiety Depression Somatization	
Behavioral Symptom Index Aggression Hyperactivity Anxiety Depression Attention Problems Atypicality Withdrawal	*Behavioral Symptom Index* Aggression Hyperactivity Anxiety Depression Attention Problems Atypicality Withdrawal	
	School Problems Attention Problems Learning Problems	

Optional Content Scales: Anger Control, Bullying, Developmental Social Disorders, Emotional Self-Control, Executive Functioning, Negative Emotionality, and Resiliency