Gender Equivalence as Perceived by Students, Parents, and Teachers on the Behavior Assessment System for Children, Second Edition

A Dissertation by

Zoë Claire Alvarez

Chapman University

Orange, California

College of Educational Studies

Submitted in partial fulfillment of the degree of

Doctorate of Philosophy in Education,

Emphasis in School Psychology

May 2015

Committee in charge:

Michael Hass, Ph.D. Chair

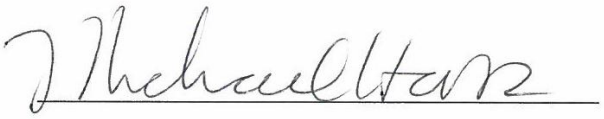
Kelly Kennedy, Ph.D.

Randy Busse, Ph.D.

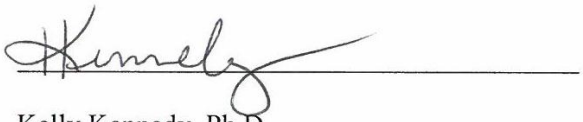
Lemuel Day, Ph.D

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

This dissertation of Zoë Claire Alvarez is approved.

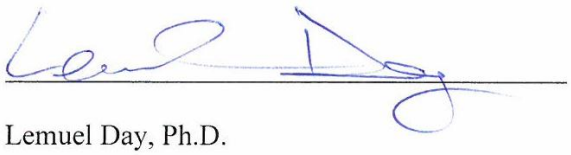


Michael Hass, Ph.D., Committee Chair



Kelly Kennedy, Ph.D.

Randy Busse, Ph.D.



Lemuel Day, Ph.D.

April 2015

Gender Equivalence as Perceived by Students, Parents, and Teachers on the Behavior Assessment System for Children, Second Edition

ACKNOWLEDGEMENTS

In dedication to pushing beyond binaries, towards equity and understanding

Words could not express my love and appreciation for my family and friends, whose undying support on my long educational journey have lit the way. Mum and Dad, without you this would not have been possible. Thank you to my loved ones for tolerating my chronic disappearance during school semesters, nourishing me at times of fatigue, and awarding me the opportunity to make you proud.

My sincerest gratitude to the faculty in the school psychology department and the College of Educational Studies, specifically Dr. Hass, Dr. Kennedy, Dr. Busse, Dr. Carriere, and Dr. Brady, who helped me develop my practitioner skills in the master's program and reignite and explore my research passions in the doctorate program.

Thank you to Rob Altmann for supporting my research ideas and providing access to the BASC-2 dataset without which this study would not have been possible.

Thank you to my dissertation committee, Dr. Hass, Dr. Kennedy, Dr. Busse, and Dr. Day, for the encouragement and direction.

VITA

EDUCATION

Chapman University	6/09-Present
Orange, CA	
Ph.D., Education	5/15 (projected)
M.A., Educational Psychology	8/11
Ed.S., School Psychology	5/12
Pupil Personnel Services Credential, School Psychology	5/12
Riverside Community College	12/08
Riverside, CA	
Early Childhood Education Assistant Teacher Certificate	
Pitzer College	5/07
Claremont, CA	
B.A., Psychology, Honors	
B.A., Gender Feminist Studies	

FELLOWSHIP, GRANTS, AWARDS

Graduate Student Scholarly/Creative Grant Chapman University	10/13
College of Educational Studies Graduate Writing Fellow Chapman University	8/10-5/12
College of Educational Studies Fellow Chapman University	2009-2010
Susan Pritzker Scholar Pitzer College	2003-2005
Fabian Núñez Scholar Pitzer College	2005-2007
Ahmanson Foundation Scholar Pitzer College	2005-2006

PROFESSIONAL EXPERIENCE

School Psychologist 3/13-Present
San Jacinto, CA
San Jacinto Unified School District

School Psychologist 8/12-12/12
Orange, CA
Rossier Educational and Mental Health Enterprises, Inc.

Classified Employee Substitute 12/05-3/13
Riverside, CA
Jurupa Unified School District

School Psychology Internship Student 8/11-5/12
Ontario, CA
Chaffey Joint Union High School District, Etiwanda High School and Chaffey High School
Riverside, CA
Jurupa Unified School District, Mira Loma Middle School and Van Buren Elementary School

School Psychology Practicum Student 8/10-6/11
Chino, CA
Chino Valley Unified School District, Chino High School and EJ Marshall Elementary School

Afterschool Tutor 8/04-3/07
Claremont, CA
Claremont Tutorial Centers

Domestic Violence Shelter Internship Student 9/05-5/06
Pomona, CA
House of Ruth

PROFESSIONAL PRESENTATION EXPERIENCE

Alvarez, Z., Mc Adams, A., & Sikkenga, A. (2014, February). *Gender and social, emotional, and behavioral referral and assessment*. Paper presented at the national convention of the National Association of School Psychologists, Washington D.C.

Alvarez, Z., Mc Adams, A. (2014, February). *School psychologists' awareness of sports related concussions*. Paper presented at the national convention of the National Association of School Psychologists, Washington D.C.

Mc Adams, A., & Alvarez, Z. (2014, February). *The sibling connection: positive interactions of siblings with a disability*. Paper presented at the national convention of the National Association of School Psychologists, Washington D.C.

Mc Adams, A., Alvarez, Z., & Baker, D. (2014, February). *Social learning for inclusive strategies in general education classrooms*. Poster presented at the national convention of the National Association of School Psychologists, Washington D.C.

Radie, J., Alvarez, Z., & Baker, D. (2014, February). *How educators can help children cope with parental deployment*. Paper presented at the national convention of the National Association of School Psychologists, Washington D.C.

Sizelove, G., & Alvarez, Z. (2013, November). *Addressing risky sexual behavior through a three-tiered model.* Paper presented at the state convention of the California Association of School Psychologists, Newport, CA.

Kennedy, K., Radie, J., Alvarez, Z., & Baker, D. (2013, November). *How educators can help children cope with parental deployment.* Paper presented at the state convention of the California Association of School Psychologists, Newport, CA.

Iranipour, A., Alvarez, Z., Riddell, C., & Weston, J. (2013, February). *Supporting Lesbian, gay, bisexual, transgender, and questioning youth in schools: A three-tiered approach*. Paper presented at the national convention of the National Association of School Psychologists, Seattle, WA.

Iranipour, A., Alvarez, Z., Padilla, A., & Weston, J. (2013, February). *Non-suicidal self injury: School-based interventions*. Poster presented at the national convention of the National Association of School Psychologists, Seattle, WA.

Alvarez, Z., Iranipour, A., Padilla, A., & Weston, J. (2012, October). *Cutting and self-injury: A behavioral approach.* Paper presented at the state convention of the California Association of School Psychologists, Costa Mesa, CA.

Alvarez, Z., Iranipour, A., Trolli, C., & Weston, J. (2012, October). *Lesbian, gay, bisexual, transgender, and questioning youth in schools: A three-tiered approach to intervention*. Paper presented at the state convention of the California Association of School Psychologists, Costa Mesa, CA.

GUEST LECTURES/INVITED ACTIVITIES

Alvarez, Z. (2014, February; 2014, October; 2015, February). *Special education and secondary transitions*. Presentation provided for school counseling graduate course at University of La Verne, Irvine, CA.

Moderator of Fireside Chat for Spring Doctoral Forum. Furlong, M. (2012, March). *What are the essential building blocks of youth psychological well-being*? Presentation and fireside chat provided for Chapman University, College of Educational Studies, Orange, CA.

Alvarez, Z., Iranipour, A., Padilla, A., & Weston, J. (2012, March). *Cutting and self-injury: A behavioral approach*. Presentation provided for school psychology graduate course at Chapman University School Psychology, Orange, CA.

PROFESSIONAL AFILLIATIONS/ACTIVITIES

Riverside Association of School Psychologists, Student Representative 8/13-Present

Chapman University Association of School Psychologists, Secretary 2011-2012

National Association of School Psychologists, Member 6/10-Present

California Association of School Psychologists, Member 6/10-Present

Psi Chi, The National Honor Society in Psychology 4/06-Present
Pitzer College Chapter Member

ABSTRACT

Gender Equivalence as Perceived by Students, Parents, and Teachers on the Behavior Assessment System for Children, Second Edition

by Zoë Claire Alvarez

Prejudice and discrimination based on gender occurs within the referral and assessment process of students' social, emotional, and behavioral functioning. Gender disproportionality has received little attention in special education research (Sullivan & Bal, 2013), yet researchers have found that males are more likely than females to be identified as having learning disability, cognitive disability, emotional disability, other health impairment, speech-language impairment, and low-incident disability (Coutinho & Oswald, 2005). Gender is often discussed as a risk factor for several mental health disorders and gender differences in the experience and prevalence of disorders are emphasized in *the Diagnostic and Statistical Manual, 5th edition* (DSM-5) (American Psychiatric Association, 2013). The purpose of this study was to examine student, parent, and teacher ratings on a behavioral rating scale, *the Behavioral Assessment System for Children, 2nd edition* (BASC-2), for similarity based on the gender of students. The normative data from the BASC-2 rating scale were examined for similarities through independent *t*- tests of equivalence.

Results indicated that the genders were similarly rated on the majority of the subscales on the BASC-2. Correlations and coefficients of determination were weak and did not demonstrate substantive strength between the BASC rating scales subscales and

gender. At the large effect size 100% equivalence was found, at the moderate effect size 86% equivalence was found, and at the small effect size 16% equivalence was found.

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

Chapter 1: Introduction

Much attention in school psychology research and school practices focuses on gender differences. Whether it is the difference between boys' and girls' literacy acquisition or classroom behavior, there appears to be a general consensus in the field of school psychology that gender implies difference. However, are these assumed differences actually important? Or, do they even exist? The current study examines if the gender of a student influences how they are perceived by those around them. Specifically, I analyzed student, parent, and teacher ratings on a popular behavior rating scale for similarities between female and male students.

In 1973, an entire issue in the journal *School Psychology Review* (SPR) was dedicated to gender issues in school psychology:

> This issue of *The School Psychology Digest* will present for your consideration some of the current thinking on sexual differences, role stereotyping, and the psychological and educational implications that sexual equality is a desideratum for both sexes a yet to be achieved goal in American society. Since the schools play a central part in maintaining this inequality, it is most important for school psychologist to be aware of these issues…School psychology possesses the vitality and the audacity to effect both attitudinal and behavioral changes in that most important social institution of the schools. What will you do to help build that nonbiased school system and to bring forth a new norm of equality? (Engin, Leppaluoto, & Fodor, 1973, pp. 2-9)

This issue was published during the second wave of feminism, fueled by the momentum of the social movement. The second wave of feminism was built on the radical feminist

protests of the women's liberation movement in Western societies pushing for empowerment for women and other disenfranchised and oppressed groups (Krolokke & Sorenson, 2006). Forty-two years later no further issues have been solely dedicated to discourse on gender roles, sexism, and gender bias. Ironically, this topic has rarely come up for discussion in school psychology since the 1973 special issue. Since then, when gender is mentioned in articles in SPR, it is as a research variable, with no attention to the concept or construct of gender.

Since the first publication of SPR in 1972, 31% of articles published referenced sex or gender, with most including sex or gender as a variable rather than broader issues of bias and inequality cited previously. Recently a series on gender was included in the National Association of School Psychologists (NASP) periodical, *the Communique*. The series included discussion on gender development, intersex and transgender issues, masculinity, and femininity. The discussion of gender, or what was formerly known as sex and sex roles, has had little attention in the field of school psychology.

Purpose of the Study

The purpose of this study was to examine student, parent, and teacher ratings on a behavioral rating scale, *the Behavioral Assessment System for Children, 2nd edition* (BASC-2), for similarity based on the gender of the students. This study is based on the gender similarities hypothesis (GSH) proposed by Hyde (2005), which states that females and males are more alike than different on most psychological traits as a framework. The normative data from the development of the BASC-2 was analyzed through descriptive statistics and equivalence testing for similarity between students' gender.

Terms and Definitions. Gender and sex were terms that were once used interchangeably. For example, in a review of the National Association of School Psychologists (NASP) publication *the School Psychology Review* the term gender was first published in an article in 1980, with the same definition as sex. For years following this gender and sex were used interchangeably in *the School Psychology Review*. However, gender has come to refer to the cultural meanings of femininity and masculinity, such as gender identity and cultural and social expectations of gender (Hawkesworth, 1997; Kimmel, 2008;). Sex connotes the biological characteristics, both anatomical and chromosomal (Kimmel, 2008). In this study I use sex to refer to biological attributes and gender to describe the social and cultural definitions of females and males.

Psychological Research: Gender Differences and Similarities

The debate about gender differences has a long history, dating back to the Ancient Greek philosophers to Darwin and Sir Francis Galton to the early figures of psychology and education, such as Thorndike and Cattell (Cattell, 1903; Epstein, 1988; Kimmel, 2008; Shields, 1975; Thorndike, 1908; Thorndike, 1910). Cattell (1903) proposed that because men were prominent figures and pioneers in fields their abilities were superior to women. Thorndike (1908) promoted the view in educational psychology that men's abilities were larger in range than women's abilities. Furthermore, education should be targeted to women's mediocrity, accepting that women's capabilities did not allow them to excel in professions (Thorndike, 1910). A counterargument was posed by Hollingworth (1914) that critiqued the perspectives that males were entitled to dominance due to a large variability of intellectual abilities. Thompson Woolley (1910) reviewed

psychology literature and concluded the mind was not a secondary sex characteristic and few to none psychological traits were biologically sex based. These arguments from the early 20th century demonstrate the length at which gender differences and similarities have been underway in the fields of psychology and education.

Trew (1998) documented other historical trends in the discussion of gender development. In the first half of the 20th century gender was seen as a socialization process by an individual's family. From the mid 1970s through early 1980s the concept of psychological androgyny was introduced; in which individuals have both masculine and feminine attributes in their behavior, attitudes, and personality (Trew). More recent trends in gender research involve social category and self-concept (Trew) and gender equality (Denmark, Klara, Baron, & Cambareri-Fernandez, 2008). These trends continue to include discourse on gender differences and similarities.

Gender Similarities Hypothesis

The gender similarities hypothesis (GSH) is based on the theory that the genders are more similar than different on most psychological traits (Hyde, 2005). This hypothesis grew out of the review of 46 meta-analyses of which sixty percent of the effect sizes for gender differences were found to have little to no effect. In a recent review of meta-analyses for gender similarities and differences Hyde (2014) found moderate gender differences in the ability to mentally rotate three-dimensional images, tender-mindedness or agreeableness, sensation seeking, interest in things rather than people, physical aggression, masturbation, pornography use, and attitudes about casual sex. Trivial to small gender differences were found in math performance, verbal ability, gregariousness, conscientiousness, reward sensitivity, negative affectivity, relational

aggression, tentative speech, oral sexual experience, attitudes regarding premarital sex and masturbation, leadership effectiveness, self-esteem, and academic self-concept (Hyde, 2014). Thus, no strong differences were found between genders in the meta-analytic reviews.

In this study I examined the gender similarities hypothesis in student, parent, and teacher ratings on the BASC-2. That is, I explored how students and those people present in students' lives on a regular basis perceive students as functioning behaviorally, emotionally, and socially in the home, school, and community settings as indicated on a behavioral rating scale. Based on the findings of Hyde (2005, 2014) I expected to find predominately similarities between behavioral ratings for female and male students and few to no differences between the sexes.

Assessment of Social, Emotional, and Behavioral Functioning

School psychologists have been described as the leading mental health experts within the school setting (National Association of School Psychologists, 2006). Specifically, as reported in *the Blueprint for Training and Practice III*, school psychologists need to be able to recognize at-risk behaviors associated with the development of mental health disorders, such as internalizing disorders, conduct disorders, and school dropout. School psychologists must be able to create preventive and intervention programs and know how to work with school staff, students, and families in addressing mental health. This involves a leadership that reduces the alienation of students and the increase of appropriate behavior through the instructional environments (National Association of School Psychologists, 2006). School psychologists are expected to be knowledgeable of mental health disorders, at risk

behaviors, and support for students and families. In addition to understanding these aspects of mental health, an important part of functioning as a mental health expert in the schools is the understanding of how gender intersects with mental health and how to work as a team to identify, assess, and support students' mental health.

Gender is often cited in *the Diagnostic and Statistical Manual, 5th edition* (DSM-5) as a risk and prevalence factor. For example, gender related diagnostic issues were reported for Disruptive Mood Dysregulation Disorder, Major Depressive Disorder, Depressive Disorder due to Another Medical Condition, Agoraphobia, Specific Phobias, Generalized Anxiety Disorder, Attention Deficit Hyperactivity Disorder (ADHD), Oppositional Defiance Disorder (ODD), and Conduct Disorder (CD) (American Psychiatric Association, 2013). For most of these disorders few studies are cited to explain gender differences, such as meta-analyses, literature reviews, and survey or interview studies, for gender related symptoms, prevalence, and disorder expression (see Appendices A-B). Thus, gender differences in the identification, prevalence, symptoms, and ultimately the treatment of mental health disorders are reported as static and factual.

Behavioral Rating Scales

Behavioral rating scales have been increasingly used in the assessment of social, emotional, and behavioral concerns since the mid-1980s (Merrell, 2008). Rating scales are forms completed by the students, parents, families, school staff, and others in the students' lives regarding their behavior, social, emotional, and adaptive skill functioning. This allows the opportunity for input on how students behave across the settings in which they spend their lives. These offer insight as to the students' skills, strengths, and problem behavior and what supports may be needed. The data gathered from rating

scales can also help to classify and diagnose and inform the development of goals, supports, and interventions (Hass, Brown, Brady, & Boehm Johnson, 2010).

Issues and cautions have been discussed in the use of rating scales. Elliott and Busse (1993) cautioned that rating scales do not capture the frequencies of behavior, yet are a snapshot of the behavior of students in various settings. Because the respondents are across settings, their ratings are situational and based on the raters' standards for behavior. Finally, the attributes of the students impact the ratings. One of the most noted attribute of students that may influence rating scale feedback is the gender of the student. With gender based expectations affecting behavioral ratings it was recommended to use separate normative group data for interpretation, comparing female to female and male to male students (Elliott & Busse). This last cautionary note of using separate gender specific normative group data for interpretation was explored in this study and informed by the results.

Method

In this study I examined the raw data from the development of the BASC-2 for similarities in student, parent, and teacher ratings between female and male students. This included reviewing the raw rating data from the preschool, child, and adolescent rating scales. The BASC-2 rating scales include survey items on problem behaviors, such as attention problems and hyperactivity related characteristics, and adaptive skills, including social skills and self-esteem. These problem behaviors and adaptive skills comprise the subscales that make up the BASC-2 rating scales (see Appendix C). To examine gender similarities on the BASC-2 rating scales the subscales of each rating scales were analyzed with tests of equivalence.

Tests of Equivalence

Tests of equivalence have been described as being appropriate statistical tests for examining similarity or equivalence rather than difference (Weber & Popova, 2012). Equivalence testing was first used for biomedical purposes to test pharmaceutical products (Schuirmann, 1987). Equivalence testing has been used to examine similarity between females' and males' performances on the Scholastic Aptitude Test math section (SAT-M) (Ball, Cribbie, & Steele, 2013). Weber and Popova (2012) promoted the use of tests of equivalence in the fields of communication and psychology to review meta-analyses. One requirement for equivalence testing is to determine the intervals or effect size guidelines by which to interpret the similarity between the groups or distributions. In this study I followed Weber and Popova's suggested interpretation recommendations for independent groups that utilize Cohen's interpretation guidelines. The equivalence testing results were interpreted based on small, moderate, and large effect sizes. Tests of equivalence were used to examine student, parent, and teacher ratings on the BASC-2 for similarities between female and male students. The following hypotheses were proposed to examine the GSH through student, parent, and teacher ratings on the BASC-2 rating scales:

Hypotheses:

1. The genders are similarly rated on *the Behavior Assessment System for Children, second edition Parent Rating Scale* (BASC-2 PRS).
2. The genders are similarly rated on *the Behavior Assessment System for Children, second edition Teacher Rating Scale* (BASC-2 TRS).

3. The genders are similarly rated on *the Behavior Assessment System for Children, second edition Self Report of Personality* (BASC-2 SRP).

Chapter 2: Literature Review

Research and discussion on gender differences can be traced through history and across disciplines. Varying fields in which the investigation of gender differences and similarities include but are not limited to psychology, evolutionary biology, sociology, and anthropology. Ancient Greek philosophers posited that males were associated with dryness and heat, and females were characterized by cold and moisture (Shields, 1975). Later, Darwin's observations and theories of man's social power and dominance due to biological and mental differences influenced theories of gender differences (Epstein, 1988; Kimmel, 2008; Shields, 1975). For example, male aggression resulting in positions of political leadership. Theories explaining the presence and existence of gender differences vary widely and the impact of these differences have been studied and debated across and within disciplinary fields. This literature review will explore the influence gender categories, roles, and stereotypes have on the practice of behavioral, social, emotional referral and assessment in the field of school psychology.

Terms within the discourse of gender differences vary as much as the theories and research on gender. The current definition of sex, as used by most social and behavioral scientists, defines it as the biological attributes of male and female, such as anatomical and chromosomal, whereas, gender refers to the cultural meanings of masculinity and femininity (Kimmel, 2008). Some scientists have argued that differentiating between the concepts of sex and gender is unacceptable because biological and social differences cannot be separated (Epstein, 1988). Hawkesworth (1997) argued that the analytic category of gender encompasses gender identity, gendered divisions of labor, gendered social relations, gender symbolism, among other topics and is in need of expanding

conceptual gender terminology leading to conceptual distinctions that allow for challenging of typical gender concepts and norms. Within this literature review, the common usage of these terms will be employed, specifically as sex refers to male and female biologically (i.e., primary and secondary sexual characteristics) and gender connotes masculinity and femininity (i.e., social and cultural expectations) (Kimmel, 2008).

Psychological Research

From early in the field of psychology there has been an abundance of research and theories attempting to explain the existence of, reasons for, and implications of gender differences. In the 19th century, Sir Francis Galton proposed the underlying assumptions of gender differences being based on female inferiority to their male counterparts (Sherif, 1979). In contrast to then contemporary psychological theory of gender differences, Thompson Woolley (1910) reviewed the psychology of gender literature and research methodology and noted "there seems to be a general trend toward the opinion that mind is probably not a secondary sexual character – in other words that there are probably few if any psychological differences of sex which are of biological origin" (p. 341). Hollingworth (1914) critiqued the belief that male achievement was attributed more so to wider male intelligence variability rather than the difference in average intelligence scores between the genders, that males were thus more entitled to dominance through their wider distribution of capabilities than females.

A prominent view in educational psychology at that time was posited by Thorndike (1908) who stated that males and females did not differ significantly in intellectual capacity, but in a primary and foundational difference that is the variability

between male and female capability. Thus, in regards to achievement, men's abilities were wider in range and men scored both higher and lower than women (Hollingworth, 1914). Men's abilities were viewed as having a larger range than women's abilities, with men having both lower and higher levels of ability overall as compared to women whose abilities were narrower in range (see Figure 1).

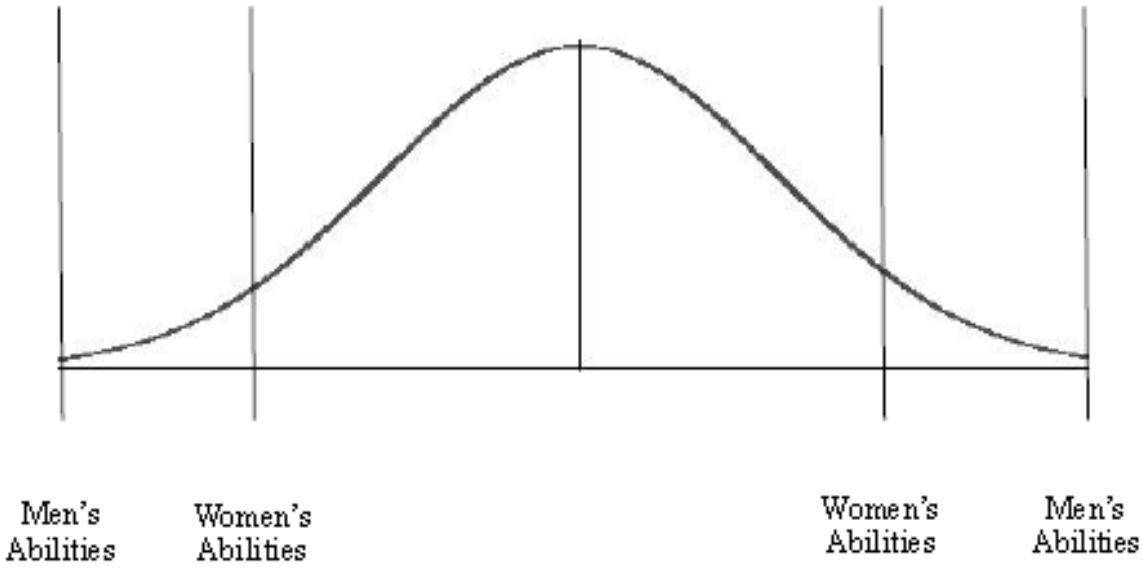


Figure 1. Men's abilities were purposed as having a larger range than women's abilities (Thorndike, 1908).

This was combined with the claim that more males were pioneers in various arenas of society and higher male achievement existed in science, technology, and politics, among other fields, to highlight female mediocrity and explain subservience. Thorndike (1908) drew on the work of Cattell (1903) who studied the accomplishments of men and concluded that, because there were more prominent historical and influential male figures than female, males were more distinguished from females and females did not differentiate from the norm. Thorndike (1910) argued that the educational system needed

to address women's restriction to mediocrity in ability and achievement in preparation for their careers in marriage, motherhood, and possible professions in nursing, architecture, teaching, or medicine. Early leaders in the study of intelligence and psychology, such as Thorndike and Cattell, examined and discussed gender differences and set the stage for the study of gender.

Much later Maccoby and Jacklin (1974) continued Woolley and Hollingworth's work on gender differences in psychological research. Maccoby and Jacklin reviewed 2,000 studies of gender differences, including those not published, and found few gender differences. Of a wide range of abilities, personality traits, cognitive abilities, and behavior, strong gender differences were found in the areas of verbal ability, visual-spatial ability, mathematical ability, and aggression. These were groundbreaking findings in the area of gender differences research. Recently, Hyde (2014) questioned the magnitude of such differences in even these areas.

Trew (1998) discussed the historical trends in interpretations of gender constructs. From the late 1890s through mid 1930s, the constructs of male and female were researched through gender differences in intelligence (Trew). This is evident in research by Thompson Woolley and Hollingworth and their examination of gender differences and critique of gender differences research (Hollingworth, 1914; Thompson, 1910). Following this period through 1974, gender research centered on individuals being socialized by their families into masculine and feminine roles (Trew, 1998). The next historical research period described by Trew spanned the mid 1970s through the early 1980s and involved the concept of androgyny. Bem (1978) emphasized the concept of a psychological androgyny, in which individuals embody both femininity and masculinity

in their identity, personality, and behavior. The most current trends in the study of gender include gender as a social category and self-concept (Trew, 1998) and gender equality (Denmark, Klara, Baron, & Cambareri-Fernandez, 2008). Research and theories within the study of gender are diverse and at times overlap in attempts to prove and explain similarities and differences between females and males.

Feminism. An important intersection between the discourse of gender similarities and differences is feminism. There are three waves of feminism most commonly discussed in Western society. The first wave of feminism occurred during the turn of the twentieth century at which point international efforts were underway igniting the "woman's rights" movement towards achieving the social and political goals for access to higher education, voting rights, and legal and moral reforms (Forestell & Moynagh, 2014, p. xvii). The second wave of feminism took place from the 1960s to 1970s and was characterized by radical protests in conjunction with other social rights movements, such as disabilities and civil rights, pushing for empowerment in oppressive political, social, and cultural systems (Krolokke & Sorenson, 2006). The subsequent third wave of feminism has been driven by the development of feminist theories and politics shaped by contradictory experiences that deconstruct categories and constructs (Krolokke & Sorenson, 2006).

Theories and Frameworks on Gender

This literature review will focus primarily on theories most applicable to gender research within the field of school psychology. Due to the wide variation in theories among gender development, researchers have discussed gender research by grouping approaches into categories (Hyde & Plant, 1995; Kimmel, 2009; Liben & Bigler, 2002).

For example, Hyde and Plant (1995) summarized perspectives on gender differences as either representing the minimalist perspective in which differences are minimal or nonexistent or the maximalist perspective, including those embodying the belief that psychological gender differences are large or are a mixture of large and small. The gender similarities hypothesis exemplifies a minimalist perspective and is based on the proposition that females and males are more similar on most psychological variables than different based on quantitative analysis (Hyde, 2005). Other types of theory classification are organized in regards to the development and recurrence of gender identity, behavior, roles, and differences. Because gender differences have been researched for an extensive time theories are abundant and they are grouped by categories to facilitate understanding.

Liben and Bigler (2002) grouped theories of gender differentiation into three categories: essentialism, environmentalism, and constructivism. Gender essentialism encompasses the theories that emphasize the role of material or physical attributes of girls and boys; an example includes evolutionary psychology. Gender environmentalist approaches rely on the environmental or societal practices that generate and maintain gender differences, such as learning theory. The third category of approaches, gender constructivism, holds the views that individuals are active agents in developing schema that influence their gender beliefs and behaviors. Social constructionist theories fall within the gender constructionist group of theories (Liben & Bigler, 2002).

Gender Essentialism

Buss and Schmitt (2011) argued a minimalist perspective of evolutionary psychology in which gender differences are expected in areas in which women and men

have confronted different adaptive problems repeatedly during the course of human evolutionary history and vice versa for gender similarities. The biopsychosocial theory incorporates biology (e.g. neurological changes in the brain in response to the environment), psychology (e.g. interests, motivation, and prior learning influence), and social (e.g. societal gender role expectations, learning, motivation) components in explaining gender differences (Halpern, 2012).

Gender Environmentalism

Social learning and social cognitive theories. In accordance to social learning theory, gender typical behaviors are performed and reinforced by others and by the individual and the rate of performance is valued differently based upon gender (Mischel, 1966). Bandura (1971) proposed that behavior is not only learned through direct observation and strengthened through immediate consequences, but influenced by perceptions and thoughts. Thus, in these types of theories people perform gender specific behaviors and are reinforced for this, which in turns increases their gender typical behavior.

Social cognitive theory emphasized that in addition to the large cognitive component of the cognitive theory of gender role development; people develop and function in broad networks with social influences (Bussey & Bandura, 1999). Societal subsystems, such as self, family, peers, education, occupation, influence the development of gender attributes and roles (Bussey & Bandura, 1999). According to social cognitive theory, a triadic reciprocity exists in which individuals exist within an interrelated system of influence, contributing to their own development and social change by means of

personal action and agency (Bussey & Bandura, 1999). Gender identity and roles are co-constructed by the individual and the social systems in which they exist.

Gender Constructivism

Social constructionism. Sociology provides a major foundational pinning of social constructionism that emphasizes the explanation of difference, power, and institutional dimensions of gender (Kimmel, 2011). Difference, based on gender, is explored through an analysis of the multiplicity, or plurality, of gender definitions. The explanation of power is undertaken by analyzing which definitions of gender become normative and who has the power to define gender. Lastly, examining the interactions between gendered institutions and individuals creates another dimension to examine gender from the social constructionism lens (Kimmel). Thus, sociology creates vantage points from which to examine the systems and the individuals in which gender is created and exists.

Feminist research is based on the philosophy that it is necessary to consider potential influencing variables within context and situations of research (Halpern, 2012). Kimmel (2008) promoted sociological assumptions that people exist within historical and cultural contexts and shape their lives based on this. Thus, in accordance with social constructionist definitions of gender and difference are dynamically influenced by power; vary across cultural institutions; vary within a single culture; vary across historical time within cultures and societal institutions; and vary over an individual's life course. From a gender constructivist perspective, people are active agents negotiating gender within societal systems and institutions defined through difference, which is ultimately driven by power.

Cognitive. Psychological theories contribute additional constructivist approaches to understanding gender. Kohlberg (1966) theorized that the development of gender specific trends follows along a cognitive trajectory, in which the individual's cognitive maturity dictates their understanding or organization of the social world, including their gender identity and gender roles. This development coincides with Piaget's stages of early cognitive development. For example, children were thought to not have a physiological understanding of gender until the concrete operations stage at approximately six to seven-years-old. Gilligan (1982) critiqued Freud's male focused theories of development and discussed development, specifically female development, in terms of social relationships and care for others. The role of the individual in social relationships influences their development.

Ecological Systems Theory

An ecological approach to understanding human development was proposed by Urie Bronfrenbrenner (1977) requiring the examination of the interactions among people and across settings and the environment. Bronfrenbrenner defined the ecology of human development as:

> The scientific study of the progressive, mutual accommodation, throughout the life span, between a growing human organism and the changing immediate environments in which it lives, as this process is affected by relations obtaining within and between these immediate settings, as well as the larger social contexts, both formal and informal, in which the settings are embedded. p. 514

Furthermore he critiqued the single setting in which experimental research commonly took place, such as a laboratory or naturalistic observation, and called for the ecological

environment or the conception of structures topographically arranged nesting within one another and each containing the next, including a microsystem, a mesosystem, an exosystem, and a macrosystem (Bronfrenbrenner, 1977). The systems progress from the physical, immediate relational context (microsystem) to the larger social, economical, political structures (macrosystem) (Bronfrenbrenner, 1977). His theory provides a framework to examine and understand social contexts and systems of human development, including behavior (Merrell, 2009) and gender.

Gender as a Variable

In psychology, gender is considered an independent variable (Sherif, 1979). Independent variables are those in experiments that are manipulated or are out of the control of the participants (Westen, 2002). Epstein (1988) posited that focusing on gender as an independent variable in explaining behavior is a conceptual and methodological bias. Constructing studies in which gender is an independent variable conveys the assumption that gender is correlated with or causes behavior, and in turn this overshadows possible behavioral similarities between the groups (Epstein, 1988). Furthermore, Epstein notes focusing on gender as an independent variable increases the difficulty of locating other variables that may contribute more than gender to behavioral differences. Sherif (1979) employed the metaphor of a railroad boxcar to describe gender as a variable, "everyone knows what it is called and what it is used for, but no one knows what is inside" (p. 101). Gender does not account for all behavior and differences found in research. Simply knowing an individual's gender does not provide insight to personality, behavior, and abilities. Furthermore, assigning gender as a research variable by which results are interpreted provides group data but no individual information.

Gender in School Psychology, Psychology, and Education

If one were to research topics on gender on the National Association of School Psychology (NASP) website (www.nasponline.org) one would find such topics as: (a) gender differences in academic skill acquisition (e.g., early literacy development); (b) gender nonconforming or gender atypical youth; (c) the ratio of female to male school psychologists in the field; (d) the lesbian, gay, bisexual, transgender, and queer (LGBTQ) community or sexual minority youth; and (e) the impact of gender in providing mental health services. In the field of school psychology, gender is overwhelmingly thought of as differences in attributes, behavior, and functioning. For example, from the search on gender on the NASP website, one of the results included a study by Friedrich, Raffaele Mendez, and Mihalas (2010) on gender-informed mental health service delivery through informing practitioners of gender differences in psychopathology prevalence, expression, and etiology. A second study reviewed gender differences in early literacy through the analysis of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) probes from kindergarten through fifth grade (Below, Skinner, Fearrington, & Sorrell, 2010). Various aspects of the profession of school psychology and gender were examined (National Association of School Psychologists, 2008; Wilson, & Reschly, 1995). The risk and needs of sexual minority and gender nonconforming youth and techniques are discussed in appropriately addressing students of the LGBTQ community (Byrne Yates, Martinez, & Harrison, 2006; Haldeman, 2000). Other topics include gender differences in cultural acceptance and career orientation (Gerner & Perry, 2000) and in gender and aggression and victimization (Murray-Close, & Crick, 2006). These research topics focus on gender differences and gender as a variable resulting in highlighting group characteristics.

To examine the reference to sex and or gender in a sample of the school psychology field literature a review was completed of all the issues of *School Psychology Review* (see Table 1). Each article published was searched for the terms sex and gender in the text. The introduction, commentary, book review, call for papers, and other supplemental pieces in volumes were also searched for the terms. Author indices were not included.

Approximately 31% of journal articles reference sex in *School Psychology Review* since being in publication and 20% of the articles reference gender as a demographic, a research category, or a research topic. During the radical second wave of feminism an entire issue was dedicated to discourse on sex roles, sexism, and practical means to address sexism and gender bias. The term gender was not used in the journal until 1980. During the 1980s the term gender grew in use. To this day sex and gender are continually used interchangeably in the journal.

Table 1
Sex and Gender in School Psychology Review

Year, Volume	Total Articles/Supplemental	Articles Referencing Sex	Articles Referencing Gender
1972, 1	48	7	0
1973, 2	42	21[a]	0
1974, 3	42	4	0
1975, 4	11	6	0
1976, 5	34	4	0
1977, 6	37	4	0
1978, 7	42	7	0
1979, 8	59	14	0
1980, 9	55	16	1
1981, 10	70	9	0
1982, 11	63	9	0
1983, 12	59	20	3
1984, 13	71	16	2
1985, 14	66	15	2
1986, 15	57	13	3
1987, 16	59	15	6
1988, 17	66	15	6
1989, 18	54	8	5
1990, 19	56	19	14
1991, 20	56	14	5
1992, 21	60	13	9
1993, 22	57	13	11
1994, 23	55	20	19
1995, 24	58	27	22
1996, 25	49	15	14
1997, 26	54	17	15
1998, 27	57	18	18
1999, 28	52	21	21
2000, 29	59	28	25
2001, 30	44	15	14
2002, 31	43	19	15
2003, 32	53	21	18
2004, 33	47	21	22
2005, 34	41	18	13
2006, 35	43	16	13
2007, 36	42	18	18
2008, 37	42	22	20
2009, 38	43	22	19
2010, 39	47	24	22
2011, 40	40	18	16
2012, 41	29	17	18
2013, 42	37	12	10
2014, 43	40	11	10
Total	2139	662	429

[a] An entire article in 1973 was dedicated to sex role related topics.

Within the last year, the NASP periodical *the Communique* featured a series on gender. The series included topics such as gender as a concept, gender identity development, theories on gender identity and development, and LGBTQ topics.

Similar to the fields of psychology and school psychology gender is also researched within educational psychology and education as a variable that is sought to explain various phenomena. A preliminary search of 'educational psychology' and 'gender' on the Discover! Search engine resulted in such topics as sex differences in math and science achievement (Reilly, Neumann, & Andrews, 2014); gender differences in self-belief, self-esteem, self-efficacy, and implicit theories of intelligence (Diseth, Meland, & Briedablik, 2014); gender differences in educational performance (Driessen & van Langen, 2013); emotional problems and school achievement (Riglin, Petrides, Frederickson, & Rice, 2014); and academic motivation and educational achievement (Vecchione, Alessandri, & Mariscano, 2014) among others.

Education and educational psychology are additional fields that encompass a wide variation on gender as researched as a variable and concept. For example, the gender discourse is reviewed as a concept within education by Lahelma (2014). Two major trends in discourses in education on gender since the 1980s have been discussed by Lahelma (2014) as being the "gender equality discourse" and the "boy discourse" (p. 171). Discourse regarding gender equity revolves around girls' and women's access to and position in education, whereas the discourse on males in education focuses on academic achievement and behavior (Lahelma, 2014). In other areas of education and psychology researchers have reviewed curriculum and pedagogy. Eaton and Rose (2013) reviewed introductory psychology textbooks for common theoretical frameworks applied

to gender based topics. Chrisler (2013) discussed the availability, relevance to professional organization mission objectives (APA Division 51: Society for the Psychological Study of Men and Masculinity), course structure, and teaching approaches for courses on the psychology of men and masculinity.

In contrast to NASP gender references online, the American Psychological Association (APA) website (www.apa.org) currently includes "Gender Issues" as a category for research. These topics are comprised of varying or different experiences of depression by men (American Psychological Association, 2005a), the impact of negative stereotypes (American Psychological Association, 2006a), and the similarities between women and men in personality, cognitive skills (American Psychological Association, 2006b), and leadership (American Psychological Association, 2005b). Whereas some topics examine gender as a variable and gender differences, APA presents more varying viewpoints on gender issues than NASP, particularly encompassing the similarities of gender in addition to differences. The research and conceptualization of gender within these areas of psychology exemplify possible beliefs of gender. Thus, this contrast in gender research scope at the national professional organization levels has implications for the fields of practice, specifically in regards to special education referral and assessment, such as providing the most appropriate interventions and services for students and their families.

Gender Prejudice and Discrimination

The National Association of School Psychologists (NASP) defines prejudice as an "attitude, opinion, or feeling formed without prior knowledge, thought, or reason" (National Association of School Psychologists, 2004). Prejudice is the source of

discrimination, and discrimination is the differential treatment that favors an individual, group, or object over others (National Association of School Psychologists, 2004). These terms were discussed in a position statement adopted by NASP in regards to race, culture, and ethnicity. In attempts to acknowledge and address the issues of racism, prejudice, and discrimination and the effects these have on students and "our nation" NASP proposes leaders in education and communities be proactive. These active roles involve promoting tolerance, pluralistic values, self-respect, and respect for others in students. School psychologists' roles include understanding how the effects of racism, prejudice, and discrimination "affect every facet of the lives of children and adults in America" (National Association of School Psychologists, 2004).

In addition to prejudice and discrimination occurring based upon people's race, culture, and ethnicity, these phenomenon take place based upon gender. Sherif (1979) defined social stereotypes as "consensual and evaluative judgments on the character, attributes, and personal qualities of individuals classified into one common social category (e.g., by national origin, race, social class, or sex) by members of another social category" (p. 123). The following study serves as an example of prejudice based on gender:, Steffens, Jelenec, and Noack (2010) found math-gender stereotypes predicted female students' academic self-concept, academic achievement, and enrollment preferences. Gender-roles and gender have been linked to varying symptoms of depression and levels of prevalence and impairment of Attention Deficit Hyperactivity Disorder (ADHD) (DuPaul, et al., 2006; Priess, Lindberg, & Hyde, 2009). Thus, prejudice and discrimination are experienced in regards to gender and as educators and

school psychologists we must understand how these impact students and ultimately our practice in the schools.

Individuals with Disabilities Education Improvement Act of 2004 and Gender

The federal legislation, the Individuals with Disabilities Education Improvement Act of 2004 (Individuals with Disabilities Education Improvement Act, 2004), requires the states to gather and submit specific data on students to the Secretary of Education. Included in these data are the number and percentage of students by race, ethnicity, limited English proficiency status, gender, and disability category who receive a Free Appropriate Public Education (FAPE); participate in regular education, in separate classrooms, schools or facilities; public or private residential facilities; the age of students between ages 14 to 21-years-old who stop receiving special education services and the reason why; and the students who are removed to an interim alternative educational setting, including long-term suspensions and expulsions, incidence and duration of disciplinary actions, hearings, and mediation (Individuals with Disabilities Education Improvement Act, 2004). Other data also required are the amount and percentage of students with disabilities by race, gender, and ethnicity who received early intervention services; those children with substantial developmental delays; and children who stopped receiving early intervention services. Disproportionality data are required to be gathered and delivered to the Secretary of Education with regards to race and ethnicity in the identification of students with disabilities; the educational settings; and the incidence, duration, and type of disciplinary actions (Individuals with Disabilities Education Improvement Act, 2004). Thus, although data are gathered to monitor the disproportionality of race and ethnic groups receiving special education, disability

categories, settings, and disciplinary actions, no data to address gender disproportionality are required. IDEIA 2004 does prohibit sex based discrimination and requires equal access to effective education (Coutinho & Oswald, 2005).

Gender and Special Education Disproportionality

States are required to provide data on the number of students receiving special education services, placement settings, disability categories, disciplinary actions, early education interventions, and the stopping of special education services in regards to disability, race, ethnicity, and gender (Individuals with Disabilities Education Improvement Act of 2004) (see Appendix D for the complete section of the act). However, technical assistance is only provided to ensure compliance and review disproportionality and overidentification in special education with regard to ethnicity and race, not gender (Individuals with Disabilities Education Improvement Act, 2004). Comments and discussion were noted regarding gender being left out of the disproportionality tracking. Responses to the comments and discussion were that adding gender to the disproportionality analysis would add "burden to the States" and no congressional intent for such analysis was noted, thus it was not felt it needed to be added to the regulations (U.S. Department of Education, n.d.).

Gender Disproportionality

Gender disproportionality has received little attention in the special education research literature (Sullivan & Bal, 2013). In a study of 24,295 students in 51 urban schools in the Midwest of the United States of America (USA) researchers found across racial groups males and students receiving free or reduced lunch were at the greatest risk for being identified as having a disability, including learning disability, cognitive

disability, emotional disability, other health impairment, speech-language impairment, and low-incidence disability.

Coutinho and Oswald (2005) collected data from 88,650 schools within 14,645 school districts on gender and special education throughout the USA. Males were more likely to be identified as having a learning disability (LD), mental retardation (MR) or an intellectual disability (ID), and a serious emotional disturbance (SED). Boys were twice as likely to be identified with LD; 1.33 times more likely to be identified with ID, and 3 ½ times more likely to be identified with SED. These odds were almost exact for the state of California. The odds ratios varied in states the most for SED from 2.2 to almost 6 times and 1.7 to 2.7 for students identified with LD. The least amount of variation was found for MR or ID. Continho and Oswald discussed that the range in ratio odds warranted the investigation of sociocultural explanations beyond biological explanations of gender disproportionality across schools and districts in the USA.

Various explanations for the gender disproportionality have been proposed and studied. Explanations focus on the overrepresentation of males and the underrepresentation of females in special education programs. Wehmeyer and Schwartz (2001) found in a records review analysis of three school districts in the South and Southwest of the USA that females were underrepresented in the LD and ID school populations. It was concluded that the underrepresentation of females in special education was a result gender bias, including behavioral expectations impacting referrals (Wehmeyer & Schwartz, 2001).

Social, Emotional, and Behavioral Functioning

Doll and Cummings (2008) argued life success is based on a foundation of emotional and social competence. The purpose of schools is to equip students with skills and knowledge to engage in productive and successful lives. To do so academic, emotional, and social competencies must be achieved. Because gender is often cited in research as a risk factor and characteristic variable in depression, ADHD, CD, and ODD (e.g., American Psychiatric Association, 2000; Bruchmüller, Margraf, & Schneider, 2012; Diamantopoulou, Verhulst, & van der Ende, 2011; Marcotte, Berger et al., 2012; Uddin et al., 2010), these disorders and research will be reviewed in reference to gender differences in prevalence, referral, and symptomology. These will be used to gain insight as to role gender stereotypes and social, emotional, and behavioral referral and assessment. These disorders have been separated into internalizing and externalizing disorders. Comorbidity occurs when individuals have more than one disorder (Satler, 2002). Students may experience both internalizing and externalizing behaviors (Merrell & Walker, 2004). Examining disorders and behaviors within the categories of internalizing and externalizing is an empirically supported and more objective approach for social, emotional, and behavioral special education referral and assessment in the schools (Merrell & Walker, 2004).

Gender and the DSM 5

Gender differences as related to mental health disorders are included in *the Diagnostic and Statistical Manual of Mental Disorders, 5th edition* (DSM 5). "Potential" gender differences are reported in the expression of mental illness (American Psychiatric Association, 2013, p. 15). As indicated in the DSM 5 gender may influence mental

illness in three ways. Gender can directly determine if an individual is at-risk for an illness, including reproductive life cycles. Examples of this are premenstrual dysphoric disorder and increased risk of major depression or mania postpartum. Gender may also moderate the risk for mental illness, which is understood through the prevalence and incidence rates for specific disorders. Lastly, gender is listed as possibly influencing the probability of specific symptoms that are experienced. An example of this would be the symptoms of Attention Deficit Hyperactivity Disorder (ADHD) that are most commonly experienced by the genders. With the exception of the physiological and biological sex related mental health illnesses, gender as it is associated with prevalence, incidence, and symptom expression characterize gender differences in the DSM 5.

Gender is promoted as being the reason or cause for differences in diagnoses and service acquisition. It was suggested in the DSM 5 that:

> Gender likely has other effects on the experience of a disorder that are indirectly relevant to psychiatric diagnosis. It may be that certain symptoms are more readily endorsed by men or women, and this contributes to differences in service provision (e.g., women may be more likely to recognize depression, bipolar, or anxiety disorder and endorse a more comprehensive lists of symptoms than men). (American Psychiatric Association, 2013, p. 15)

Information on the genders is provided to address gender at different levels. One level is at the diagnostic criteria level at which gender-specific symptoms are included. The next level of information includes gender-related specifiers that provide information on gender and diagnosis. The final level of information on gender is provided in sections for pertinent issues to gender and diagnoses considerations titled *Gender-Related Diagnostic*

Issues (American Psychiatric Association, 2013). Providing this section for disorders facilitate the location of gender differences that have been observed in clinical and community settings.

The studies cited for gender related prevalence and diagnostic concerns were reviewed and summarized (see Appendices A-B). An electronic version of the DSM-5 contained in-text citation for studies, however a hard copy of the DSM-5 contained no citation information for mental health disorders. In both texts there are often statements with no direct or obvious citation to indicate a source. Most cited studies were based on epidemiological surveys, interviews, literature reviews, and meta-analyses that included adult samples, and were based on self-reported information. There were five citations regarding Social Anxiety Disorder (Social Phobia), the disorder with the most citiations. This lack of systematic and consistent source citation makes it difficult to interpret and analyze the studies credited with finding and documenting gender related symptoms and specifiers for mental health disorders.

The mental health disorders as defined and categorized in *the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, Text Revised* (DSM-IV-TR) have a larger research base and will thus be the primary focus of the current literature review with some discussion of the gender related information included for selected disorders in *the Diagnostic and Statistical Manual of Mental Disorders, 5th edition* (DSM-5) when available.

Internalizing disorders. Merrell (2009) defined internalizing problems as including symptoms and characteristics associated with depression, anxiety, social withdrawal, and somatic complaints. These problems are often difficult to detect because

they are over-controlled and internally directed attributes, usually characterized by emotional problems. Much research has been conducted on depressive and anxiety disorder and gender.

Depression and anxiety. According to the DSM-IV-TR depressive disorders include Major Depressive Disorder, Dysthymic Disorder, and Depressive Disorder Not Otherwise Specified (American Psychiatric Association, 2000). These depressive disorders include some level of depressed mood, loss of interest or pleasure, weight loss or gain, insomnia or hypersomnia, loss of energy, feelings of worthlessness or guilt, and inability to concentrate, among other symptoms. Noted gender differences in Major Depressive Episode include females being at higher risk for developing a Major Depressive Episode during their lives than males, a differentiated risk emerges during puberty and in adolescence, and twice as many females have depressive episodes twice as frequently as males. Boys and girls are equally affected by Major Depressive Disorder, but adolescent and adult females were twice as likely to be affected as adolescent and adult males. Similarly, Dysthymic Disorder occurs equally in girls and boys, but is two to three times more likely in women than men.

In the DSM-5 depressive disorders now include Disruptive Mood Dysregulation Disorder, Major Depressive Disorder, Persistent Depressive Disorder (Dysthymia), Premenstrual Dysphoric Disorder, Substance/Medication-Induced Depressive Disorder, Depressive Disorder Due to Another Medical Condition, Other Specified Depressive Disorder, Unspecified Depressive Disorder, and Specifiers for Depressive Disorders (American Psychiatric Association, 2013). "Gender-Related Diagnostic Issues" are noted for Disruptive Mood Dysregulation Disorder, Major Depressive Disorder, and Depressive

Disorder Due to Another Medical Condition. Two studies were cited in which males were found to have higher rates of Disruptive Mood Dysregulation Disorder than females and it is asserted that the difference in prevalence rates helps to differentiate between this disorder and Bipolar Disorder of which the genders have equal prevalence rates (see Appendix A). Women continue to be noted as having a higher prevalence rate for Major Depressive Disorder, be at greater risk for suicide attempts, and at lower suicide completion rates than men; however, women and men have similar symptoms, course, treatment responses, and functional impacts. Lastly, gender-related diagnostic concerns associated with Depressive Disorder Due to Another Medical Condition are those gender differences that are found in the medical condition itself rather than the depressive disorder.

There are numerous anxiety disorders, including Panic Disorder, Agoraphobia, Specific Phobia, Social Phobia, Generalized Anxiety Disorder, and various other forms of anxiety disorders in the DSM-IV-TR (American Psychiatric Association, 2000). Panic Attacks can include heart palpitations, feelings of choking, dizziness, sweating, fear of dying and losing control. Agoraphobia is anxiety due to being in crowded places or situations in which escape may be difficult. A Specific Phobia is an irrational, marked, and persistent fear of an object or situation. Social Phobia is a fear of social situations and unfamiliar people. Generalized Anxiety Disorder is characterized by excessive worry and anxiety and physical symptoms. Panic Disorder without Agoraphobia is diagnosed twice as often in women than men and Panic Disorder with Agoraphobia is three times as often diagnosed in women than men. In total, Agoraphobia is diagnosed more often in women than men. The ratio of women to men diagnosed with Specific

Phobias is two to one. In clinical studies, females and males are equally diagnosed with Social Phobia, while in epidemiological and community samples women were found to have Social Phobia more than men. This was also the case with Generalized Anxiety Disorder; the rates of diagnosis in males and females varied from clinical to community settings with females outnumbering males in both samples.

The amount of anxiety disorders in the DSM-5 has grown to include Separation Anxiety Disorder, Selective Mutism, Specific Phobia, Social Anxiety Disorder (Social Phobia), Panic Disorder, Panic Attack Specifier, Agoraphobia, Generalized Anxiety Disorder, Substance/Medication-Induced Anxiety Disorder, Anxiety Disorder Due to Another Medical Condition, Other Specified Anxiety Disorder, and Unspecified Anxiety Disorder (American Psychiatric Association, 2013). In the DSM-5 one study was cited in reporting that girls experience more reluctance to attend school than boys and that males indirectly express more separation fear (e.g. being away from home alone, when partners do things without them, or when they cannot be in contact with their partners) as related to Separation Anxiety Disorder (Allen et al., 2010). For Social Anxiety Disorder, or Social Phobia, males are more likely to experience paruresis, or the inability to urinate when others are around, and two studies were cited as indicating comorbidity differences in the genders. Panic Disorder was not indicated as having any features that varied by gender, but a possible genetic difference associated with Panic Disorder was purported in two studies (see Appendix A). Panic attacks are reported in the DSM-5 as being more common in females than males with no variation in features or symptoms. Differing comorbidity was reported for Agoraphobia for females and males, including mental disorders for females and substance abuse disorders for males. In clinical and

epidemiological studies females were diagnosed more often with Generalized Anxiety Disorder and comorbidity patterns varied for the gender, but symptoms were similar for males and females.

Gender and Prevalence. Prevalence rates of depressive and anxiety disorders in females and males have been examined. Rushton, Focier, and Schectman (2002) analyzed the National Longitudinal Study of Adolescent Health survey (N=13,568) and found 30% of adolescents reported depressive symptoms and 10% reported moderate to severe symptoms. Gender was the sole socio-demographic consistently associated with depression over time, although during initial ratings females, older adolescents, and ethnic minorities were more likely to report depressive symptoms (Rushton, Focier, & Schectman, 2002). In a review of 796 four-year-old children, Lavigne et al. (2009) found no gender differences for generalized anxiety disorder, major depressive disorder, dysthymia, or separation anxiety disorder.

The age of onset of depressive and anxious symptoms has been another component of psychological research. In a cross-cultural examination, Wade, Cairney, and Pevalin (2002) analyzed data from Canada, Great Britain, and the United States (N=16,168) and found females had higher rates of depression with a pattern of onset at 14 years of age. Another estimate highlights the age frame of 15 through 18 years of age as a critical time for the development of depressive symptoms as doubling in females as compared to males (Hankin et al., 1998).

Gender and symptom expression. In a study of 177 6th through 8th grade students, Dixon, Scheidegger, and McWhirter (2009) found males reported higher levels of depression and anxiety than females. Gender differences were examined in genetic and

environmental determinants of depressive symptoms and adolescent depression (Uddin et al., 2010). A genotype was found that manifests as a protective factor against depression and does so differently for males and females; mainly in males the genotype manifestation is modified by the social environment.

Chaplin, Gillham, and Seligman (2009) found physiological anxiety symptoms were equally predictive of future depression symptoms for girls and boys. However, symptoms of worry, oversensitivity and overall anxiety symptoms were more predictive of depression symptoms for girls.

Marcotte, Alain, and Gosselin (1999) found more females reported depressive symptoms than males. Studies have shown men are less likely to report depressive symptoms than women (Berger et al., 2012; Kessler et al., 1993). Berger et al. (2012) found men were less likely to do so when the label of depression was used in association to factors that are outside of individual control. In a clinical sample, Newman, Fuqua, Gray and Simpson (2006) found anger and depression occurred equally in the men and women who self-referred for mental health services. As supported by their research, Newman, Fuqua, Gray, and Simpson cautioned that when people seek mental health services women do not necessarily outnumber men, as is usually found in general population studies. Furthermore, the depressive symptoms of women were no more intense than men's symptoms. Because of this, it is necessary for clinicians to "reflect on their own expectation for gender differences in the presentation of depression" (Newman, Fuqua, Gray, & Simpson, 2006; p. 160). This can be accomplished through an idiographic approach, or an individualized approach, to exploring depression and anger with each client.

Explanations for gender differences in depression. The gender intensification hypothesis has been used to explain the higher rates of depression in female adolescents than males (Hill & Lynch, 1983; Hyde, Mezulis, & Abramson, 2008; Marcotte, Alain, & Gosselin, 1999). Hill and Lynch (1983) posited that during early adolescence and puberty onset individuals are pressured to conform to gender roles, behavior, and attitudes. Males are encouraged to adopt instrumental characteristics, such as assertiveness and independence; while females are socialized to be expressive and compliant. Marcotte, Alain, and Gosselin (1999) tested the gender intensification hypothesis and found the possession of instrumental characteristics was negatively correlated with depression, specifically in males, while the expressive characteristic was not correlated with depression. Thus, these results partially supported the gender intensification hypothesis.

Externalizing disorders. Merrell (2009) posited that in contrast to internalizing disorders, externalizing problems are undercontrolled and outwardly directed concerns, primarily associated with behavioral problems rather than emotional. Symptoms and characteristics of externalizing disorders include overactive, acting-out, aggression, antisocial, and disruptive behavior.

Attention deficit hyperactivity disorder (ADHD). The American Psychiatric Association (2000) defined ADHD as a persistent pattern of hyperactivity and/or inattention that is more severe than same-aged peers and exhibited frequently. There are three subtypes of ADHD, ADHD Combined Type, ADHD Predominantly Inattentive Type, and ADHD Predominantly Hyperactive-Impulsive Type. These disorders are more

prevalent in male than females depending on the subtype and clinic versus community settings.

In the DSM-5 the description of ADHD now refers to the persistent pattern of inattention and/or hyperactivity/impulsivity that impairs functioning or development (American Psychiatric Association, 2013). Similar to the DSM-IV-TR, the DSM-5 includes the gender differences that males are more likely to have ADHD in childhood, 2:1, and adulthood, 1.6:1 in the general population and females are more likely to display inattentive characteristics. However, less than half of the studies reviewed discussed gender in the findings (Polanczyk et al., 2007; see Appendix B).

Prevalence study findings. Bauermeister et al. (2007) found ADHD was 2.3 times more likely to be diagnosed in boys than girls and boys were more likely to be suspended from school than girls. Bruchmüller, Margraf, and Schneider (2012) found among a sample of 473 German child and adolescent therapists, therapists were more likely to diagnose boys with ADHD than girls. Results from the National Health Interview Survey indicated parents reported ADHD symptoms that were clinically significant in 4.19% of males and 1.77% of females (Cuffe, Moore, & McKeown, 2005).

Gender and symptom expression. A cross-cultural review of ADHD found higher ratings of boys displaying inattention and hyperactivity-impulsivity than girls (Bauermeister, Canino, Polanczyk, & Rohde, 2010). Similar to these findings, Lavigne et al. (2009) found boys were more likely to be diagnosed with ADHD-Inattentive Type. Diamantopoulou, Henricsson, and Rydell (2005) found when boys exhibited moderate levels of ADHD symptoms peers were more tolerant than when these were displayed by girls. Thus, it appeared ADHD symptoms were incompatible with the feminine

stereotype, causing girls to be exposed to greater difficulties in peer relations as compared to boys.

Another component of gender and ADHD that has been researched is impairment. Rucklidge (2006) found females with ADHD exhibited similar impairments in neuropsychological functioning as males with ADHD, including intellectual functioning and executive functioning. One gender difference presented during the study was males with ADHD showed higher levels of impairment in inhibition when compared to females with ADHD. Graetz, Sawyer, and Baghurst (2005) found girls with ADHD were more likely to express somatic complaints and boys displayed more impaired school functioning. No gender differences were found in externalizing behavior in the study.

In a meta-analysis of 38 studies, Gershon (2002) found males were reported as displaying more impairment in hyperactivity, inattention, and impulsivity than females. Females were rated by teachers and parents as being less hyperactive and inattentive than males. Teachers rated females as displaying fewer externalizing symptoms than males. These results differed from other meta-analytic reviews of ADHD and gender.

A second meta-analysis revealed gender differences in ADHD impairment varied based on non-referred and clinically referred sample groups, including non-referred females displaying more impairment in inattention, internalizing behavior, aggression, and peer dislike (Gaub & Carlson, 1997). Within the clinically referred group, the sole gender difference was males demonstrated less impairment in inattention than females. Overall, behavioral gender differences between females and males were comprised of females exhibiting lower levels of hyperactivity, conduct disorder diagnoses, externalizing behaviors, and greater intellectual impairment.

Comorbidity studies on ADHD have included reviewing gender differences and similarities. Levy, Hay, Bennett, and McStephen (2004) found no gender differences in the tendency for comorbidity of females and males with ADHD and externalizing disorders, such as Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD). Females were rated as having symptoms of Seasonal Affective Disorder (SAD), a depressive disorder, associated with the inattentive type and GAD with the combined type (Levy et al., 2004). From this, researchers postulated that the severity of ADHD symptoms accounts for comorbidity rather than gender. Bauermeister et al. (2007) found boys with ADHD Combined Type were more likely to have comorbidity with mood disorders than girls, and girls with ADHD Inattentive Type were more likely than boys to have comorbidity with anxiety disorders.

Referrals. Researchers have examined the referrals from parents and teachers of children with behavioral concerns, specifically in terms of ADHD. Mothers and teachers rated boys as displaying higher scores of ADHD and ODD than girls (Waschbusch & King, 2006). Coles, Slavec, Bernstein, and Baroni (2010) found teachers rated girls as displaying significantly more impairment and being in need of services more than boys. These results were gathered from 50 elementary school teachers who responded to vignettes and rating skills examining student impairment. Girls were more likely to be referred for services when displaying ADHD and ODD symptoms rather than inattentive symptoms.

Sciutto, Nolfi, and Bluhm (2004) found teachers were more likely to refer male students who exhibited hyperactivity, gender differences in referral for the other categories of symptoms were not statistically significant. Researchers posited if gender

bias in teacher referral was present it existed in response to hyperactivity and not milder or more severe symptoms. Sciutto, Nolfi, and Bluhm explained their findings as evidence that the gender differences in teacher referrals mirrored the gender differences in symptoms expression and teacher bias influence referral practices and decisions. Reid et al. (2000) found evidence of gender differences in teachers' perceptions of symptom severity on behavior ratings, suggesting rather than perceptions of gender differences being limited to symptoms, gender differences are broad in nature. Additionally, gender differences were more pronounced as symptom severity increased.

DuPaul et al. (2006) found females and males with ADHD displayed the same level of impairment in school functioning, below that of their peers in the social, behavioral, and academic domains. Gender differences that were found were teacher perceptions of student academic, social, and behavioral functioning. Teachers reported female students demonstrated higher levels of academic motivation and study skills than male students. Male students were reported by teachers as displaying more severe symptoms of ADHD than female students. When females with ADHD were compared to females without ADHD, teachers perceived them as exhibiting more severe symptoms, however, when males with ADHD were compared to males without there was no such discrepancy. Within the areas of social, emotional, and behavioral functioning, females were reported by their teachers as displaying more symptoms of internalizing and externalizing problems, school problems, inattention, and low adaptive skills.

Hypothesis for gender differences in ADHD. Ohan and Visser (2009) tested the disruptive behavior hypothesis through the distribution of vignettes to 96 parents of children with ADHD symptoms and 140 elementary school teachers. The disruptive

behavior hypothesis holds that those students who are most disruptive behaviorally are referred for services or assessment. What was found did not support the disruptive behavior hypothesis, but rather disruptiveness was a factor for parents and teachers in seeking help for female students with ADHD but not male students with ADHD. Teacher and parents were more likely to refer male students for services, despite disruptive behavior. These findings were attributed to the possibilities of perceptions that female students' disruptive behavior was not viewed as acceptable and male students were more likely to benefit from services.

Oppositional Defiance Disorder (ODD) and Conduct Disorder (CD). The DSM-5 describes Oppositional Defiance Disorder (ODD) as a recurring pattern of disobedient, negative, hostile, and defiant behavior toward authority figures. This might include the tendency to lose one's temper, defy or refuse to comply with rules or requests by adults, become easily annoyed, blame others, and be spiteful. Up until puberty, ODD is more prevalent in males, after puberty rates equal out in the genders. Symptoms are similar in males and females with the exception of males being more confrontational and experiencing symptoms that are more persistent. Studies report inconsistent and insignificant findings for ODD and gender differences (Boylan et al., 2007; Nock et al., 2007; see Appendix B). In the DSM-5 no gender related section differences or issues were included.

Conduct Disorder (CD) is defined as a persistent and repetitive behavioral pattern that is characterized by the violation of others' basic rights, rules, and major age-appropriate social norms. Four main groups of CD include aggressive conduct, nonaggressive conduct, deceitfulness or theft, and serious rule violations. CD is more

common in males than females. Conduct problems for individuals with CD vary by gender. Males are more likely to display fighting, stealing, vandalism, school discipline difficulties, and confrontational aggression. Whereas females with CD exhibit behaviors such as truancy, lying, running away, substance use, prostitution, and nonconfrontational behavior. In the DSM-5 the gender related differences in behaviors commonly exhibited by the genders are still reported and supported by only one study. Additionally, with the citation of another single study, females were indicated as tending to participate in more relational aggression, whereas males partake in both relational and physical aggression toward others.

Prevalence. Research on conduct disorders (i.e. ODD, CD) has been primarily conducted with males (Lehto-Salo, Närhi, Ahonen, & Marttunen, 2009). Lavigne et al. (2009) found in early childhood there were no gender differences in the prevalence of ODD and impairment.

Diamantopoulou, Verhulst, and van der Ende (2011) found no gender differences in the development from early childhood to adolescence in symptoms of ODD and CD. There were no associations between CD and ODD, rather CD symptoms best predicted future CD symptoms. As referenced previously, mothers and teachers rated boys as displaying higher scores of ODD and ADHD than girls (Waschbusch & King, 2006).

Gender and symptom expression. In a study of parent and teacher reports of ODD, Munkvold, Lundervold, and Manger (2009) found with the exception of emotional symptoms, boys were rated by teachers as displaying more symptoms of ODD than girls. No gender differences were found in the parent ratings of children. Overall, children reported as displaying ODD symptoms had an increased risk for being rated as displaying

emotional symptoms, hyperactivity or inattention, and peer problems when compared to children who were not identified as exhibiting ODD symptoms.

Rydell (2010) examined environmental factors and demographics in relation to ODD and ADHD. The results indicated mothers with less than nine years of education, single or step-parent families, and those of non-European descent were variables associated with high symptoms levels of ADHD and ODD. Few gender effects were found, one of which boys were found to be more vulnerable to family stress than girls as related to the expression of ADHD and ODD symptoms.

Offord et al. (1996) reported the importance of the consideration of sources of information in classifying CD and ODD. Lehto-Salo, Närhi, Ahonen, and Marttunen (2009) found no gender differences in the number of CD and aggressive symptoms. Adolescent females were more commonly reported as experiencing paternal domestic violence and relationship problems, major depression, anxiety disorders, substance abuse disorders, and suicidal behavior as compared with males. Results indicated boys from poor families were more likely to be identified as having ODD by teachers. Children identified as having ODD and CD by teachers were likely to have a depressed parent and live in a "dysfunctional family" (Offord et al., 1996, p. 1084).

Parent and teacher referrals. Offord et al. (1996) examined parent-teacher agreement in report of CD and ODD symptomology using three separate data integration strategies. The percentages of agreement between parent and teacher report of students having CD and ODD were low. Munkvold, Lundervold, Lie, and Manger (2009) found teachers and parents reported boys as displaying more symptoms of ODD than girls, with teachers reporting so more often. Frequencies at which ODD symptoms were reported as

occurring varied by informant and gender of student. Few students were identified by both parents and teachers as having ODD.

Hypothesis for gender differences in ODD and CD. According to the gender paradox hypothesis, girls with ODD experience more co-occurring mental health symptoms than boys with ODD (Munkvold, Lundervold, & Manger, 2011). The gender paradox, also known as group resistance and the threshold effect, is based on the reasoning that females experience more impairment from anti-social behavior because they display a higher threshold of risks and symptoms when compared to males (Moffitt, Caspi, Rutter, & Silva, 2001). Moffitt, Caspi, Rutter, and Silva (2001) examined the threshold effect comparing boys' and girls' environmental, abilities, and behavior with and without CD. No support for the threshold or gender paradox hypothesis was found in the study. Munkvold, Lundervold, and Manger (2011) studied parent and teacher response of co-occurring mental health symptoms of 7,007 children ages seven to nine with and without ODD. The results of this study did not support the gender paradox hypothesis.

Gender as a Group Descriptive

Researchers purport study findings as confirming and refuting gender differences and similarities. Theories are formed, partially supported, or not supported, spurring the creation of further theories. Contradicting research results obscure the discourse on gender similarity and difference. The extensive history of this ongoing debate conveys the improbability of a singular concluding theory. The extent to which these studies vary exemplifies the diverse perspectives of gender differences and similarities characteristic of the discourse. During the end of the 19th century, Wilhelm Windelband, a philosopher,

reasserted the role philosophy plays in scientific inquiry (;Mos, 1998;). In 1894, Windelband separated scientific thought into two categories, nomothetic or idiographic (as cited in Lamiell, 1998). Windelband's definition of idiographic is more inclusive than merely the study of individuals, and encompasses sought-after knowledge of groups or what could be, thus being historically and culturally situated (as cited in Lamiell, 1998, p.27). Nomothetic knowledge, research or discussion, refers to aggregates, or groups, empirically and statistically studied, rather than all people (Lamiell, 1998). In reference to Windelband, research including gender as a variable is nomographic, or statistical research of groups which attempts to apply rules of behavior to groups (i.e., gender) (as cited in Lamiell, 1998).

When addressing social, emotional, and behavioral concerns, individual characteristics and needs outweigh group descriptive characteristics. Assessment and intervention address one student's needs, adhering to the typical behavior of a group in place of individualized assessment and service is discriminatory, unethical, and ineffective (Merrell, Ervin, & Gimpel, 2006; National Association of School Psychologists, 2004). Within the field of school psychology, best practice approaches to assessment inherently address gender stereotypes and prejudice that occur.

Assessing Social, Emotional, and Behavioral Functioning

Should gender differences be taken into consideration when assessing students? This question comes into play when determining whether to compare deviation or disorder based on combined or separate gender norms or on an absolute standard or a relative standard (Reid et al., 2000). An absolute standard translates to a single standard or threshold for all groups, across possible mitigating factors, requiring the pooling of

varying scores from different groups (Reid et al., 2000). In contrast, a relative standard varies according to groups, thus thresholds are based on the context and characteristics of the group (Reid et al., 2000). Reid et al. (2000) found evidence for using relative standards or separate norms in assessing for ADHD, supported by group differences between females and males, and European Americans and African Americans. Waschbusch and King (2006) found when using gender specific norms, a subset of female students were rated with higher than average scores in ADHD and ODD yet were not diagnosed using the *DSM-IV* criteria, suggesting the *DSM-IV* diagnostic criteria may under identify girls for further assessment and possible intervention (Waschbusch & King, 2006).

Merrell, Ervin, and Gimpel (2006) called for movement beyond stereotyping groups and individuals (e.g., based on cultural and linguistic diversity) through the understanding of within- and between-group differences and the possible implications and relevance of these. Group differences include how individuals who are members of a group differ from those of other groups, whereas individual differences are the various ways in which individuals differ from one another. When understanding individual and group differences it is important to keep in mind two considerations: (a) significant variability exists *within* groups and (b) significant overlaps exist *between* groups (Merrell, 2003; Merell, Ervin, & Gimpel, 2006). Thus, although there are descriptive characteristics of a group as a whole (e.g., gender), those characteristics do not describe all individual members of the group (Merrell, Ervin, & Gimpel, 2006). In the field, school psychologists must be cognizant of this and refrain from drawing inferences of

students' characteristics, abilities, behavior, and functioning based on group membership (Merrell, Ervin, & Gimpel, 2006).

In their study of gender and the relationship between anger and depression, Newman, Fuqua, Gray, and Simpson (2006) found no gender differences. These findings supported their recommendation for clinicians to be cognizant of efforts to base clinical judgments and decisions based on observations of individual clients instead of gender stereotypes and prevalence patterns. Children's environments influence their mental health, as seen in the association between demographic risk factors and ADHD and ODD symptoms (Rydell, 2010). Whereas it is important to be aware of possible group characteristics (e.g., higher prevalence of depression for adolescent girls), individual attributes and needs should be paramount. Adhering to group characteristics, in this case gender stereotypes, may cause the oversight of individual characteristics.

Paradigm shift in school psychology. Currently, within the field of school psychology a paradigm shift is occurring in the philosophical foundations of school psychology practices (Reschly, 2008). NASP has supported this shift deeming the problem-solving model best practice in place of the correlational or refer-to-test model (Thomas, & Grimes, 2008). Reschly (2008) described the correlational model as an emphasis in assessment and study of natural cognitive, social, emotional, and physical variations in people. Relationships between these variations and functioning or performance are studied and outcomes are thought to be alterable through environment change. The experimental model incorporates problem-solving in comparing the effectiveness of interventions seeking causal information, not simply correlation and relationship data. Although problem solving has been discussed as being relatively new

to school psychologists, it has been embedded in psychological and educational research since the 1950s and the basic tenants date back farther to the origins of humans (Reschly, 2008). Problem-solving is an individualized that is approach practical for working with students, families, and school staff.

Data-driven problem solving model. The National Association of School Psychologists has issued a position statement in support of the problem-solving model approach to provide services to students in inclusive settings (National Association of School Psychologists, 2009). Merrell, Ervin, and Gimpel (2006) credited Susan Gray with the promotion of the problem-solving approach for school psychologists in 1963. Reschly (2008) traced the problem-solving steps to the beginnings of human history, noting the basic steps of implementing a solution, observing effects, and continuing or changing a practice to influence the outcome. Detailed steps of the problem-solving model include (a) problem identification; (b) problem analysis; (c) develop and implement intervention; and (d) evaluate intervention and follow-up (Merrell, Ervin, & Gimpel, 2006). These steps are cyclical and intended to be undertaken as often or refined as necessary. Identifying the problem is one of the initial steps in the model, requiring a definition of the problem and validation of its presence and need to be addressed. Data must be gathered once the problem is agreed upon to seek to answer the question of why the problem occurs. Then possible solutions are generated for the problem and put into place. These interventions are evaluated and follow-up action is taken as appropriate, which might involve redefining the problem, collecting more data, or implementing a different intervention. Frequent monitoring of progress and altering of ineffective interventions creates a self-correcting feature of the problem-solving model (Reschly,

2008). Within the field of education, specifically school psychology, the problem-solving model is outcome focused and context specific. Direct methods of data collection are utilized in gathering information that is objective, observable, and measurable (Merrell, Ervin, & Gimpel, 2006).

The problem solving model is an approach ideal for special education referral and assessment that is individualized to each student. Working as a team to identify strengths, pin point problems, generate interventions, and monitor progress specific to the student is a way to focus on the student and avoid potential gender bias. Instead of basing expectations on the gender of the student the problem solving model creates a forum in which to understand a student's current behavioral concerns in comparison and contrast to their typical behavioral functioning, rather than on cultural or societal expectations based on gender. Nondiscriminatory assessment is another best practice that inherently avoids gender bias and is complimentary to the problem-solving model and is promoted by NASP.

Nondiscriminatory assessment. Ortiz (2008) proposed that nondiscriminatory assessment is "simply good assessment" because this type of assessment applies to all individuals in whom diversity spans beyond race, culture, ethnicity, and language (p. 661). Nondiscriminatory assessment is a process for evaluation to be completed in the least discriminatory manner as possible for every individual; this type of assessment is not limited to those who are different in varying ways. Nondiscriminatory assessment does not rely on a single unbiased test instrument, but it is "an outline for systematic evaluation that promotes equity and justice across and throughout the entire process and not on techniques specific to one population or another" (Ortiz, 2008, p. 661). Much like

the problem solving model, the assessment process begins with a hypothesis focusing on the external or environmental causes for student's difficulties, rather than attributing these as being intrinsic.

Conclusion

Many theories from various disciplines propose explanations for gender differences and similarities. Ecological system theory has been most relevant with the best practices of school psychology (Merrell, 2009). Regardless of the dispute of the presence of gender differences, referral and assessment of students must be individualized and occur from the problem-solving approach. It is important to keep stereotypes in mind when addressing referral concerns and throughout assessment. Nondiscriminatory assessment practices can be applied when working with all students, as is appropriate in reducing gender-based assumptions and stereotypes. As mental health professionals in the school, school psychologists must recognize and understand internalizing and externalizing behaviors in conjunction with prevention and intervention programs designed to address such problems (National Association of School Psychologists, 2006). This includes being cognizant of the gender variables associated with risk and prevalence research and the potential gender stereotypes that affect the referral, assessment process, and potentially treatment.

Future researchers might examine school psychologists' views and beliefs of gender in reference to social, emotional, and behavioral functioning of students. Gender-based stereotypes may be foci of future studies that include the identification of internalizing and externalizing behaviors by school psychologists. Finally, the use of

nondiscriminatory assessment and the problem-solving approach by school psychologists should be surveyed to examine practices in school settings.

Chapter 3: Methodology

Hyde (2005) proposed the gender similarities hypothesis, positing that females and males are more similar than different on most psychological variables based on quantitative analysis. This hypothesis was born out of the analysis of 128 meta-analyses conducted on psychological gender differences where sixty percent of the effect sizes for gender differences were found to be small. This led Hyde to propose that females and males are more alike than different on most psychological variables.

Ball, Cribbie, and Steele (2013) explored how the GSH might be further examined through a statistical analysis of the performance of men and women on *the Scholastic Aptitude Test Math* (SAT-M) test. Tests of equivalence were used. These involved two simultaneous one-tailed t-tests. Other statistics utilized included Cohen's *d* effect size for correlations, the correlation coefficient *r*, the coefficient of determination r^2 (the proportion of variance in SAT-M scores explained by gender), and the coefficient of nondetermination or overlapping between scores λ. The purpose of the study was to examine gender similarities as rated by students, parents, and teachers on a behavior rating scale. Thus, employing tests of equivalence will allow the framing of the study and discussion of results to be centered on similarities of groups rather than differences (Ball, Cribbie, & Steele, 2013). This also expands the language for interpreting statistical results of similarity as opposed to the rejection of a null hypothesis and implied or default similarity of groups (Ball, Cribbie, & Steele, 2013).

Behavioral Rating Scales

Rating scales are used to examine how students, family members, guardians, school staff, or others who interact regularly with students, perceive a student's behavior or skills in various areas of functioning, such as social and emotional domains. Hass, Brown, Brady, and Boehm Johnson (2010) discussed several reasons why rating scales are important components of behavior assessments: they summarize observation data across settings and time; are used within a problem solving model to classify and diagnose; scale normative data is used for problem identification and diagnostic purposes; and the data can inform the development of interventions, goals, and supports for the student. Rating scale use in evaluating social, emotional, and behavioral functioning and skills has increased considerably since the mid-1980s (Merrell, 2008). Common uses of rating scales include screening, progress monitoring tools, and assessing children's social, emotional, and behavioral functioning. Because of the prevalent use of rating scales it is important to examine possible implications on the perception of problem behaviors and competencies of students by the students themselves and the individuals present most in their lives.

Elliott and Busse (1993) discussed five issues with the use and interpretation of behavior rating scales. Ratings are snapshots or summaries of specific behaviors that occur in varying frequencies. This variance in the frequency of behavior is not captured by rating scales. Responses are also the judgments of the raters and are situational, vary by environment, and are based on the raters' standards for behavior. Elliott and Busse defined social validity as the intended domains of behavior on a scale that are relevant to the functioning of the student as rated by significant individuals in their lives. Lastly, the

attributes of the children influence the rating of their behavior, with the most noted attribute being sex. Because of this it is often proposed that separate sex normative group data be used for interpretation (Elliott & Busse, 1993). Reynolds and Kamphaus (2004) cautioned that the use of gender-specific norms when using the BASC-2 ratings scales as consideration is needed on the type of behaviors that are being examined and how gender might relate to such behaviors.

Measurement Instrument

The Behavior Assessment System for Children, Second Edition (BASC-2)

The BASC-2 is the most commonly used broad-band rating scale by school-based practitioners (Hass et al., 2010). The BASC-2 was selected because it is an omnibus instrument that attempts to examine adaptive skills, problems behaviors, and school problems to aide in the differential diagnosis or educational classification and intervention development. During the test construction process of the BASC-2 gender, or sex differences as referred to in the BASC-2 manual, were analyzed. Furthermore, separate or combined norm sex groups are provided for interpretation options when using the BASC-2 with the conclusion that some sex differences exist in specific areas, which will be discussed in greater detail. To examine gender similarities of students' self, parent, and teacher ratings on the BASC-2 sets of normative data will be examined. The child form of the BASC-2 *Parent Rating Scale* (PRS) and *Teacher Rating Scale* (TRS) include ages 6 to 11-years-old from parent and teacher respondents. The adolescent rating scale includes students 12 to 21-years-old and self, parent, and teacher respondents. Reviewing both the child and adolescent rating scales will provide insight into the perception of social, emotional, and behavioral functioning of students.

The *Behavior Assessment System for Children, Second Edition* (BASC-2) is a system for the evaluation of children and young adult's behavior and self-perceptions incorporating multiple data collection methods and dimensions (Reynolds & Kamphaus, 2004). This assessment tool was created to assist in the differential diagnosis, educational classification, and development of treatment programs of and for children and young adults. The BASC-2 is used with youth from two to 25-years-old. Its multimethod design includes rating scales for the individual (Self-Report of Personality (SRP)), parents (Parent Rating Scale (PRS)), and teachers (Teacher Rating Scale (TRS)); a Structured Developmental History (SDH); and a classroom observation form for recording and classifying behavior (Student Observation System (SOS)). Because the BASC-2 is used to measure both positive and adaptive behaviors and negative or clinical concerns it has been described as multidimensional assessment system. Merrell (2009) described the BASC-2 as representing the best available behavior rating scale with an impressive empirical research base and few drawbacks (i.e., the length of the rating scales). The TRS-C, TRS-A, PRS-C, and PRS-A are the strongest components of the BASC-2 (Merrell, 2009).

During the development of the BASC-2, attempts were made to closely represent the United States (US) population in 2001 (Reynolds & Kamphaus, 2004). Data collection took place in 40 states, 257 cities, and over 375 sites in the north east, north central, south, and west of the US. Over 13,000 TRS, PRS, and SRP were included in the General norm from public and private schools, mental health clinics and hospitals, preschools, and daycares. The racial or ethnic groups identified were White, Hispanic, African American, and other. Students in general education classrooms were sampled for

the General norms and students from special education programs and clinics or treatment centers for the Clinical norms. The young adult samples were collected from colleges for the General norms and special education programs for the Clinical norm groups. Attempts were made to match the demographics of the US population as closely as possible for the General norms, including sex, race or ethnicity, mother's education, and geographic region. Data collection site coordinators and sites were compensated for participation and parents and teachers were paid for completed forms.

As reported in the BASC-2 manual, the first step in developing the General norms for the rating scales was to determine if sex and age differences were large enough to warrant separate norm subgroups (Reynolds & Kamphaus, 2004). Normative age groups were then divided into 2 through 3-years-old, 4 through 5-years-old, 6 through 7-years-old, 8 through 11-years-old, 12 through 14-years-old, and 15 through 18-years-old. Sex differences were then explored through the comparison of mean scores for females and males.

Reynolds and Kamphaus (2004) concluded that statistically significant sex differences were consistent on many of the BASC-2 scales and composites; yet few reached a half of the standard deviation. Noted sex differences were that females were rated on the PRS and TRS higher than males on Adaptability, Social Skills, and Functional Communication. Males were rated on the PRS and TRS as higher on Hyperactivity, Aggression, Atypicality, and Attention Problems. Females reported higher scores on the SRP on Anxiety and Somatization. On the SRP males rated themselves as higher on Attitude to School, Attitude to Teachers, and Sensation Seeking.

The measurement authors caution that separate sex norms are offered due to these differences, but separate norms should not be used for interpretation in all cases.

Teacher Rating Scales (TRS): Reliability and Validity

Reliability or the consistency of a test (Salkind, 2008) of the TRS was addressed during test construction through internal consistency, test-retest reliability, and interrater reliability (Reynolds & Kamphaus, 2004). Internal consistency refers to how similar, or consistently, the scale or items measures a concept. A coefficient alpha statistic was calculated to determine the internal consistency of scales and composites on the TRS for the age based normative groups, sexes, and clinical groups (LD, ADHD). The lowest coefficient alphas found on the general norm samples ranged from .74 to .79 at the preschool and 6 to 7-years-old male, female, and combined sex normative groups for the Anxiety scale; .77 for preschool ages on the Somatization scale for both sexes and combined normative groups; .77 on the Atypicality and .78 on Withdrawal for females ages 6 to 7-years-old; .77 on the male and combined sex ages 2-3 normative groups for Withdrawal; and .76 for males and .79 for combined sexes ages 2 to 3-years-old on the Adaptability scale (see Table 2).

Table 2
TRS Internal Consistency Reliabilities of the General Norm Groups

	Preschool					
	Ages 2-3			Ages 4-5		
Scale	Combined	Female	Male	Combined	Female	Male
Adaptability	.79		.76			
Anxiety	.75	.76	.74			.78
Somatization	.77	.77	.77			.78
Withdrawal	.79		.77			
	Child					
	Ages 6-7			Ages 8-11		
Scale	Combined	Female	Male	Combined	Female	Male
Anxiety	.78	.79	.77			
Atypicality		.77				

On the clinical norm samples most coefficient alpha reliabilities were above .80. Those below .80 included Somatization with .73 for males, .79 for females, and .76 for combined on the All Clinical preschool age norm group; .70 for female and .78 for combined sexes on the Somatization scale child Attention Deficit Hyperactivity Disorder (ADHD) group; Learning Problems scale coefficients included .78 for combined sexes, .78 for female, and .79 for male for Learning Disability (LD) child groups, .79 for combined sexes for ADHD Child group, .76 for female adolescent LD and ADHD groups, and .75 for ages 19 to 21-years-old combined sexes; .79 for the Withdrawal LD adolescent female group; .76 for the combined sexes and .71 male LD child groups and .79 for the female adolescent group on the Leadership scale; and .78 for the female adolescent LD group on the Functional Communication normative group.

Test-retest reliability refers to the agreement between the same respondent at different points in time (Reynolds & Kamphaus, 2004). Teachers within each age level of the TRS completed the rating scale twice with a time span of 8 to 65 days between the administrations. Sample sizes included 32 females and 37 males for preschool; 40

females and 43 males for child, and 48 females and 40 males for adolescent. The correlations were adjusted for one "that better reflects what one would expect if a much larger sample were used" (Reynolds & Kamphaus, p. 134). Test-retest reliabilities were above .80 for those subscales not provided (see Table 3). Those scales with the highest test-retest reliabilities were Hyperactivity, Aggression, and Attention Problems and those with the lowest were Anxiety and Somatization.

Table 3
TRS Test-Retest Reliabilities

	Preschool		Child		Adolescent	
Scale	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability
Anxiety	.78	.77	.73	.76	.66	.64
Depression					.76	.64
Somatization	.72	.72	.65	.74	.74	.79
Withdrawal					.71	.74
Adaptability		.78			.78	
Social Skills	.77	.76			.74	.74
Leadership					.78	
Functional Communication					.79	

Reynolds and Kamphaus (2004) discussed four types of validity evidence explored and provided for the TRS. The first type of validity discussed occurred during test construction and the alignment of the BASC-2 rating scales to diagnostic systems. During this part of development of the BASC-2 rating scales expert professional input and children, parent, and teacher perceptions were included. A second type of validity was calculated through scale intercorrelations and factor analysis of scale and composite groupings, including covariance structure analysis and principal-axis factor analysis. Hyperactivity, Aggression, Conduct Problems, and Attention Problems had the highest correlations on the TRS. Anxiety, Depression, and Somatization had generally weak correlations. Depression showed moderately high correlations with Hyperactivity,

Aggression, and Conduct Problems. Somatization had the weakest correlations to all other scales. The adaptive scales had higher correlations with one another than the clinical scales had with one another. The third validity type involved examining correlations between the scale and composite scores with those obtained from other behavior measures. Other behavior measures that were compared with the PRS were *the Achenbach System of Empirically Based Assessment* (ASEBA) *Caregiver-Teacher Report Form*, *the Conners' Teacher Rating Scale-Revised,* and the original BASC TRS. Correlations between the TRS Child (TRS-C) and ASEBA Teacher's Report were completed on a sample size of 57 and the TRS Adolescent (TRS-A) on a sample size of 39. Correlations were moderate to moderately high for Anxiety, Depression, and Somatization ranging from .63 to .77 on the TRS-C and ranged from low to adequate .22 to .78 on the TRS-A. Hyperactivity, Aggression, and Conduct Problems correlations were higher ranging from .67 to .80 on the TRS-C and .57 to .89 on the TRS-A. Similar patterns were found with *the Conners' Teacher Rating Scale-Revised.* Correlations between the BASC TRS were all above .80 with the lowest correlations on the TRS-C (.82) and TRS-A (.80) Conduct Problems. The final type of validity was drawn from examining the scoring profiles of children grouped by clinical diagnoses or educational classifications. Group score profiles were created for children with diagnoses of ADHD, Bipolar Disorder, Depression Disorders, Emotional/Behavioral Disturbance, Hearing Impairment, LD, Mental Retardation or Developmental Delay, Motor Impairment, Pervasive Developmental Disorders (PDD) (including Autism and Asperger's Disorder), and Speech or Language Disorder. Most profile groups were relatively similar with the exception of the Bipolar Disorder and Depression Disorders profile groups, however, the

sample sizes for this group was small with 7 students in the TRS-C Bipolar Disorder group, 11 students in the TRS-A Bipolar Disorder group, and 9 students in the Depression TRS-A.

Parent Rating Scales (PRS): Reliability and Validity

Internal consistency estimates were lower for the PRS than the TRS. Coefficient alpha reliabilities were found in the .70s on the general normative groups in Hyperactivity: Adolescent; Aggression: Preschool; Anxiety: Preschool; Depression: Preschool; Somatization: Preschool, Child, Adolescent; Atypicality: Preschool, Adolescent; Withdrawal: Child, Adolescent, and Activities of Daily Living: Preschool, Child, and Adolescent. In the clinical normative groups reliabilities in the .70s were found on Anxiety: Adolescent; Depression: Preschool; Somatization: Preschool; Atypicality: Adolescent; Withdrawal: Adolescent; Attention Problems: Child; Leadership: Child; and Activities of Daily Living: Child and Adolescent (see Table 4).

Table 4
PRS Internal Consistency Reliabilities of the General Norm Groups

	Preschool					
	Ages 2-3			Ages 4-5		
Scale	Combined	Female	Male	Combined	Female	Male
Hyperactivity						
Aggression	.78	.72				
Anxiety	.77		.73			
Depression	.78	.79	.77			
Somatization	.79	.77		.79		.79
Atypicality	.77	.72		.75	.74	.75
Withdrawal						
Activities of Daily Living	.77	.78	.76	.70	.73	.65

	Child					
	Ages 6-7			Ages 8-11		
Scale	Combined	Female	Male	Combined	Female	Male
Hyperactivity						
Aggression						
Anxiety						
Depression						
Somatization			.79	.79	.79	
Atypicality						
Withdrawal	.77	.78	.76			
Activities of Daily Living	.73	.73	.73	.76	.76	.76

	Adolescent					
	Ages 12-14			Ages 15-18		
Scale	Combined	Female	Male	Combined	Female	Male
Hyperactivity				.79	.76	
Aggression						
Anxiety						
Depression						
Somatization			.77			
Atypicality	.79	.77		.79	.79	.79
Withdrawal				.79	.79	.79
Activities of Daily Living	.76	.74	.75	.72	.70	.73

Parents or caregivers within each age level of the PRS completed the rating scale twice with a time span of 9 to 70 days between the administrations. Sample sizes

included 48 females and 39 males for preschool; 39 females and 38 males for child, and 51 females and 37 males for adolescent. Again, correlations were adjusted to correct for a larger sample size. Test-retest reliabilities were above .80 for the Hyperactivity preschool and child, Aggression child, Anxiety adolescent, Depression child and adolescent, Somatization preschool and adolescent, Atypicality child, Withdrawal preschool and child, Attention Problems child and adolescent, Adaptability child, and Social Skills child scales. Some of the lowest test-retest reliabilities were on the Anxiety, Depression, Somatization, Attention Problems and Social Skills scales. Activities of Daily Living had some of the highest test-retest reliabilities of all the scales (see Table 5).

Table 5
PRS Test-Retest Reliabilities

	Preschool		Child		Adolescent	
Scale	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability
Hyperactivity					.74	.75
Aggression	.74	.75			.72	.78
Anxiety	.71	.78	.65	.73		
Depression	.66	.73				
Somatization			.66	.65		
Atypicality	.72	.73				
Withdrawal					.78	.75
Attention Problems	.70	.72				
Adaptability	.77	.74			.74	.75
Social Skills	.74	.72			.77	.82

Interrater reliability, the agreement of ratings between two different raters at the same point in time, was calculated for the PRS. Each child was rated between 0 to 70 days by two parents or caregivers with a sample size of 40 for the PRS- Preschool (PRS-P), 43 for the PRS-Child (PRS-C), and 51 for the PRS-Adolescent (PRS-A). Most interrater reliabilities were in the .70s with the highest reliabilities on the Hyperactivity

(PRS-P), Anxiety (PRS-C), Depression (PRS-A), Atypicality (PRS-P), Withdrawal (PRS-A), Attention Problems (PRS-C and PRS-A), Adaptability (PRS-C and PRS-A), Activities of Daily Living (PRS-C and PRS-A), and Functional Communication (PRS-P, PRS-C, and PRS-A) (see Table 6).

Table 6
PRS Interrater Reliabilities

	Preschool		Child		Adolescent	
Scale	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability
Hyperactivity	.81	.74	.78	.73	.74	.69
Aggression	.59	.53	.58	.53	.79	.76
Conduct Problems			.65	.61	.79	.80
Anxiety	.56	.57			.69	.66
Depression	.71	.59	.77	.67	.86	.78
Somatization	.71	.70	.58	.53	.67	.55
Atypicality	.82	.73	.71	.67	.86	.78
Withdrawal	.79	.78	.70	.55	.81	.79
Attention Problems	.78	.77	.80	.75	.80	.79
Adaptability	.74	.78	.82	.78	.78	.80
Social Skills	.64	.64	.75	.70	.72	.71
Activities of Daily Living	.86	.78	.80	.76		
Functional Communication			.82	.75	.82	.76

The same four types of validity discussed for the TRS were included for the PRS. The scale intercorrelations for the PRS included small positive correlations found on the PRS-P and PRS-C for the Anxiety with Social Skills, Activities of Daily Living, and Functional Communication. The highest intercorrelations were found on Hyperactivity, Aggression, Conduct Problems, Attention Problems, and Atypicality scales, and the lowest intercorrelations were found on between Anxiety and Somatization and the other scales. Depression was moderately intercorrelated with other clinical scales. Finally,

similar to the TRS, the adaptive scales had higher correlations with one another than did the clinical scales.

Behavior measures the PRS was correlated with included *the Achenbach System of Empirically Based Assessment* (ASEBA) *Child Behavior Checklist*, *the Conners' Parent Rating Scale-Revised, the Behavior Rating Inventory of Executive Functioning* (BRIEF), and the original BASC PRS. Correlations between the PRS Preschool (PRS-P) and ASEBA Child Behavior Checklist were completed on a sample size of 53, the PRS Child (PRS-C) on a sample size of 63 to 65, and the PRS Adolescent (PRS-A) on a sample size of 67. Correlations were low to moderately high on all the PRS-P ranging from .32 (Raw reliability between Anxiety and Anxious/Depressed) to .79 (Raw reliability between Hyperactivity and ADHD). Correlations on the PRS-C were also low to moderately high ranging from .38 (Adjusted reliability between Depression and Withdrawn/Depression) to .77 (Adjusted reliability between Aggression and Aggressive Behavior). The PRS-A also had many low to moderately high correlations with a range from .30 (Raw reliability between Anxiety and Anxiety Problems) to .77 (Raw and adjusted between Aggression and Aggressive Behavior).

Correlations calculated between the PRS-C and *the Conners' Parent Rating Scale-Revised* were higher and ranged from .40 (raw reliability between Anxiety and Anxious-Shy) to .84 (Raw and adjusted reliability between Aggression and Oppositional). On the PRS-A correlations with *the Conners' Parent Rating Scale-Revised* ranged from .35 (Raw and adjusted reliability between Anxiety and Anxious-Shy) to .68 (Raw reliability between Hyperactivity and DSM-IV: Hyperactivity/Impulsivity). Between the BRIEF and the PRS-C correlations ranged from

.11 (Raw and adjusted reliability between Working Memory and Anxiety) to .77 (Raw reliability between Hyperactivity and Behavioral Regulation Index) and -.77 (Raw reliability between Adaptability and Shift; raw reliability between Adaptability and Global Executive Composite). Correlations between the BRIEF and PRS-A ranged from .02 (Adjusted reliability on the Adaptability and Monitor) to .83 (Raw reliability between Hyperactivity and Global Executive Composite). Lastly, the BASC-2 PRS were correlated with the BASC PRS. The lowest correlations were found between the PRS-P on Atypicality and Attention Problems and on the PRS-C on Atypicality, all other correlations were above .80. Similar findings were indicated on the PRS as the TRS by the analysis of the profiles of clinical groups. Bipolar or depressive disorders had elevated profiles, yet small sample sizes of Bipolar Disorder PRS-C 8, Bipolar Disorder PRS-A eight, and Depression Disorders PRS-A 11. The profiles for the PDD varied with the highest scores on the Atypicality, Withdrawal, Adaptability, Social Skills, and Functional Communication scales.

Self Report of Personality (SRP) Scales: Reliability and Validity

Internal consistency estimates were below .80 on most subscales on the SRP. Coefficient alpha reliabilities were found in the .70s or below on the general normative groups in Attitude to Teachers: Child and Adolescent; Sensation Seeking: Adolescent; Locus of Control: Child and Adolescent; Sense of Inadequacy: Child and Adolescent; Somatization: Adolescent; Attention Problems: Child and Adolescent; Hyperactivity: Child and Adolescent; Relations with Parents: Child; Interpersonal Relations: Adolescent; Self-Esteem: Child and Adolescent; and Self-Reliance: Child and Adolescent. Clinical normative group reliabilities were found to be in the .70s on child

scale only for Social Stress and Relations with Parents and on the adolescent scale only for Sensation Seeking and Somatization. Coefficient alpha reliabilities were also found to be in the .70s for both child and adolescent groups on Attitude to School, Attitude to Teachers, Atypicality, Locus of Control, Sense of Inadequacy, Attention Problems, Hyperactivity, Interpersonal Relations, Self-Esteem, and Self-Reliance. Lower coefficient alpha reliabilities were found in the .60s on Attitude to Teachers: Child; Sensation Seeking: Adolescent; Somatization: Adolescent; Attention Problems: Child and Adolescent; and Self-Reliance: Child and Adolescent (see Table 7).

Table 7
SRP Internal Consistency Reliabilities of the General Norm Groups

	Child		
	Ages 8-11		
Scale	Combined	Female	Male
Attitude to Teachers	.72	.71	.72
Locus Control	.76	.76	.76
Sense of Inadequacy	.78	.79	.78
Attention Problems	.76	.75	.76
Hyperactivity	.76	.74	.76
Relations with Parents		.79	
Self-Esteem	.77	.77	.76
Self-Reliance	.71	.71	.71

	Adolescent					
	Ages 12-14			Ages 15-18		
Scale	Combined	Female	Male	Combined	Female	Male
Attitude to Teachers				.79	.77	
Sensation Seeking	.69	.68	.67	.70	.68	.69
Locus of Control				.78	.77	.79
Sense of Inadequacy		.79		.79	.78	.79
Somatization	.67	.70	.61	.67	.68	.63
Attention Problems	.7	.79	.78	.79		.77
Hyperactivity	.76	.77	.74	.74	.73	.76
Interpersonal Relations	.79	.79		.78	.77	.78
Self-Esteem			.78			
Self-Reliance	.68	.69	.68	.70	.71	.71

Students at the child and adolescent levels were selected to complete the same rating scale form within 13 to 66 days between administrations. Sample sizes included

58 females and 55 males for child and 73 females and 34 males for adolescent. . Like the PRS and TRS, the SRP test-retest reliability estimates were adjusted for a restricted range. Because of the variation in test-retest time periods it is difficult to determine how strong the reliabilities estimates are for subscales. Test-retest reliabilities of subscales were in the .80s for Attitude to School: Child and Adolescent; Depression: Adolescent; Attention Problems: Adolescent; and Relations with Parents: Adolescent. The rest of the test-retest reliabilities were in the .70s or below (see Table 8).

Table 8
SRP Test-Retest Reliabilities

	Child		Adolescent	
Scale	Raw Reliability	Adjusted Reliability	Raw Reliability	Adjusted Reliability
Attitude to School	.66	.68	.70	.73
Atypicality	.75	.75	.75	.79
Locus of Control	.77	.75	.72	.74
Social Stress	.67	.63	.74	.74
Anxiety	.73	.72	.69	.70
Depression	.74	.71		
Sense of Inadequacy	.72	.72	.72	.74
Somatization			.71	.67
Attention Problems	.68	.70		
Hyperactivity	.70	.71	.68	.69
Relations with Parents	.64	.63		
Interpersonal Relations	.75	.71	.75	.75
Self-Esteem	.73	.67	.78	.78
Self-Reliance	.64	.64	.63	.61

Three types of validity were discussed for the SRP. These included the scale intercorrelations and factor analysis; correlations with other behavior scales; and the

score profiles of students with specific clinical diagnoses or educational classifications. On the child scale the highest intercorrelations were found on Attitude to School, Attitude to Teachers, Locus of Control, Social Stress, Anxiety, Depression, and Sense of Inadequacy; while the lowest intercorrelations were found on Attitude to School and Self-Reliance. Unlike the PRS and TRS, the SRP-C had higher intercorrelations of the clinical subscales than the adaptive subscales. Intercorrelations on the adolescent SRP (SRP-A) were highest for Attitude to School, Attitude to Teachers, Attention Problems, Hyperactivity, Depression, and Sense of Inadequacy. The lowest intercorrelations were found on the SRP-A for Sensation Seeking, Self-Reliance, and Relations with Parents. Similar to the SRP-C, the SRP-A had stronger intercorrelations for the clinical subscales than for the adaptive subscales.

Behavior measures the SRP was correlated with included *the Achenbach System of Empirically Based Assessment* (ASEBA) *Youth Self Report Form*; *the Conners-Wells' Adolescent Self Report Scale*; *the Children's Depression Inventory* (CDI); and *the Revised Children's Manifest Anxiety Scale*. Correlations between the SRP-A and ASEBA Youth Self-Report were completed on a sample size of 51 adolescents ages 12 through 18-years-old. Correlations were ranged from -.01 (Raw and adjusted between Self-Reliance and Oppositional Defiant Problems) to .84 (Raw between Anxious/Depressed to Anxiety). The strongest correlations between the SRP-A and *the Conners-Wells' Adolescent Self Report Scale* were .64 (Raw between DSM-IV Hyperactive Impulsive and Hyperactivity). The correlation between the Total CDI score and Depression on the SRP-C was .29, but was higher at .69 between the Total CDI score and the SRP-A. The correlations between the SRP-C and *the Revised Children's Manifest Anxiety Scale* was .60 and .49 between the SRP-A and *the Revised Children's Manifest Anxiety Scale.* Unlike the PRS and TRS, the SRP clinical group analysis revealed few elevated profiles for the clinical groups. Those groups that were elevated, Bipolar Disorder, were slightly elevated on the Attention Problems and Hyperactivity subscales. Like the PRS, the sample size for Bipolar Disorder on the SRP was eight students.

Psychometric Properties of the BASC-2 Rating Scales

Internal consistency estimates were higher for the PRS and TRS than the SRP. The lowest internal consistency estimates were found on the SRP adolescent Somatization, Sensation Seeking, and Self-Reliance several of which were in the .60s. Whereas all the TRS and PRS subscales internal consistency estimates were .70 or above.

Test-retest reliabilities were based on ratings that were gathered during large time frames. For example the test-retest time frame for the TRS was 8 to 65 days; for the PRS it was 9 to 70 days; and for the SRP it was 13 to 66 days. The time frames were not consistent across the rating scales. Thus, there are temporal issues in the test retest reliability estimates and it is difficult to determine how strong the reliabilities are.

Content validity was addressed in the development of the BASC-2 rating scales by attempting to align the rating scales to diagnostic systems. Content validity refers to the extent to which items on a test represent the behavior or strengths purported to being measured. Input was gathered from expert, students, parents, and teachers. Factor analyses were completed. The rating scales were also correlated with narrow band measures to ensure the concepts or behaviors being measured. However, various subscale concepts are not included in diagnostic systems, such as the DSM-5, or are measured by narrow band measures. Examples of these include Sensation Seeking, Leadership, and Functional Communication among others.

The response format for most of the BASC-2 rating scales are subjective response options. With the exception of the SRP *true* or *false* items, the majority of the items on the rating scales require respondents to select *never*, *sometimes*, *often*, or *almost always*. These responses are not defined or described on the BASC-2 and thus are left open to interpretation for the respondents. One rater's definition of *never* will vary from another's definition of *never.* These definitions might also vary by behavior and setting for the rater. The BASC-2 rating scales are perception based measures directly influenced by the perceptions of the raters.

Sample and Participant Selection

The participants for the current study were drawn from the standardization sample of *the Behavior Assessment System for Children, second edition* (BASC-2) (Reynolds & Kamphaus, 2004). As discussed previously, the total standardization sample for the BASC-2 consisted of more than 13,000 TRS, PRS, and SRP cases between the ages of 2 to 18 years and 706 SRP cases between ages 18 to 25 years. Raw data used from the BASC-2 development were de-identified and participants were anonymous. The BASC-2 manual provides evidence that the General norm samples closely resembles the 2001 census estimates in most areas, including parental education, race/ethnicity, geographic region, and special education classification. The most variation in percentages were at ages 2 and 18, at which ages fewer cases were collected. All standard scores in the present study are derived from the national standardization sample from the BASC-2 standardization and norm development.

Sample size and power. The power of a statistical test is the probability of obtaining sample results that lead to rejecting the null hypothesis when it is false; that is a difference or association beyond chance is found to exist (Coladarci, Cobb, Minium, & Clarke, 2008; Jaccard & Becker, 2002). This is commonly referred to as power. Sample size has been noted as being one of the most important factors affecting power (Coladarci et al., 2008). To an extent the larger the sample size the more powerful the test is (Coladarciet al., 2008; Jaccard & Becker, 2002); however, there is a point of diminishing value at which the test is sufficiently powerful to confidently draw conclusions (Jaccard & Becker, 2002). Thus, at that point and as the sample size increases so does the probability of finding small but possibly meaningless mean differences as being

statistically significant (Cribbie, Gruman, & Arpin-Cribbie, 2004). Additionally the probability of committing a Type II error increases with very large sample sizes (Coladarci et al., 2008). Type II error is the failure to reject the null hypothesis when it is false (Jaccard & Becker, 2002). Thus, given the size of the sample of the current study there is an increased probability of Type II error and finding potentially meaningless statistically significant mean differences.

Study Variables

A total of 26 index ratings from the SRP, PRS, and TRS were included in the study. Gender, or sex, as indicated on the BASC-2 form is a second variable. The ratings from the standardization of the BASC-2 SRP, PRS, and TRS were analyzed in the form of raw scores.

Independent variable. In the current study gender is the independent variable. During the standardization of the BASC-2 gender was equally represented across age groups and rating scales (Reynolds & Kamphaus, 2004).

Dependent variables. The self, parent and teacher ratings from the BASC-2 are the dependent variables. These include response on 26 index scales. Definitions of the scales and descriptions are found in Appendix C. The scales vary by age and respondent. The PRS and TRS scales include Activities of Daily Living, Adaptability, Aggression, Anxiety, Attention Problems, Atypicality, Conduct Problems, Depression, Functional Communication, Hyperactivity, Leadership, Learning Problems, Social Skills, Somatization, Study Skills, and Withdrawal. The SRP scales are Anxiety, Attention Problems, Attitude to School, Attitude to Teachers, Atypicality, Depression, Hyperactivity, Interpersonal Relations, Locus of Control, Relations with Parents, Self-

Esteem, Self-Reliance, Sensation Seeking, Sense of Inadequacy, Social Stress, and Somatization.

Data Analysis

Data gathered from the BASC-2 children, parent, and teacher ratings were organized by age and rating scale. The rating scales were then reviewed for gender similarities by age and respondent. Similarities in gender were examined through tests of equivalence and the review of effect sizes. Ball, Cribbie, and Steele (2013) proposed the use of equivalence tests to test the gender similarities hypothesis. Hyde (2005) developed the gender similarities hypothesis which stated that there are more similarities between the two genders on psychological traits than there are differences. With this in mind I examined how students, parents, and teachers rated students on the BASC-2, specifically if there are more similarities in how students were rated based on gender than differences and how strong these similarities and/or differences were. This was based on a procedure proposed by Schuirmann (1987) using two one-tailed *t* tests.

Tests of equivalence or equivalence testing. Tests of equivalence, or equivalence testing, have been used in the biomedical and communication fields (Schuirmann, 1987; Weber & Popova, 2012), but not within the field of education or school psychology. Schuirmann (1987) proposed using two one-sided tests in place of a null hypothesis significance test approach to assess the equivalence of bioavailability or bioequivalence studies for pharmaceutical product tests. These types of tests are appropriate when the purpose of the study is to demonstrate equivalence, or similarity, rather than difference (Weber & Popova, 2012). Null hypothesis testing has been discussed in the statistical literature for years with little attention on methodological

alternatives (Schuirmann, 1987). The two one-sided test approach breaks down null hypothesis testing into two one-sided hypotheses to determine an equivalence interval between the groups. Thus, more flexibility in how the hypotheses for tests of equivalence are framed is offered by these types of tests (Weber & Popova, 2012). In these hypotheses, μ represents the mean rating responses on the SRP, PRS, and TRS scales and scale items for female and male students and *D* is the equivalence interval. Thus, in accordance with Cribbie, Gruman, and Arpin-Cribbie (2004) the composite hypotheses can be expressed as:

$$H_{01}: \mu_F - \mu_m < D$$

$$H_{02}: \mu_F - \mu_M > -D$$

Thus, if both hypotheses are rejected an interval in which the groups are similar can be determined. It also implies that $\mu_F - \mu_m$ lies within the range of (*-D, D*) and the means are considered equivalent (Cribbie, Gruman, & Arpin-Cribbie, 2004). This approach is identical to a confidence interval approach in that equivalence is declared if the equivalence interval contains the confidence interval 1-2α (α=alpha).

Determining equivalence intervals or maximum no effect. Conducting tests of equivalence requires the determination of comparison equivalence intervals or of a minimum substantial effect. Two approaches to conducting tests of equivalence exist in the literature. These are similar in purpose yet vary in the methods to compare similarity and disprove significant difference.

The approach utilized by Ball, Cribbie, and Steele (2013) relied upon the creation of three equivalence intervals to compare. These were arbitrarily set at 1/3, 2/3, and one SD, with SD representing the average of female and male standard deviations.

Determining the intervals included consideration of the effect sizes and the percentages of overlap in the distribution and each gender at the top and bottom 5% of the distribution. Two cautionary points are made by Ball et al. in determining the intervals. First determining or defining what a meaningful difference is quantified as can be difficult. The second point involved the need to reflect on what is the largest difference between the means of the population that would be meaningless (Ball, Cribbie, and Steele, 2013).

Weber and Popova (2012) proposed performing equivalence testing based on determining the minimum substantial effect or maximum no effect (Δ). In this approach the maximum no-effect level is determined and written into the hypothesis:

$$H_0: \left|\text{effect}\right| \geq \left|\Delta\right| \qquad H_1: \left|\text{effect}\right| < \left|\Delta\right|$$

The aim of these tests is to prove similarity with no effect or difference in either direction. To determine the maximum no effect Weber and Popova encouraged reviewing Δ effect sizes based on theory and existing research in the field and topic. Thus the selection of Δ effect sizes was be deemed as not completely arbitrary, but based on previous research and theoretical and practical considerations, contrary to the claims by Ball et al. (2013).

In the current study effect sizes were not available on the rating scales in the research literature. A search of Discover, Eric EBSCO, and PsychINFO did not yield information regarding effect sizes for the BASC SRP, PRS, and TRS. Thus, the approach proposed by Weber and Popova (2012) was employed to interpret the data based on Cohen's *d* guidelines for the minimum substantial effect.

Assumptions. Schuirmann (1987) discussed the assumptions of the two one-sided test when comparing it to the power approach. These include the assumptions of a normal distribution of data; equal within subject variances of the groups; and an equal number of observations or participants in each group for simplified comparisons.

To examine the normality of the rating subscale distributions the minimum and maximum raw ratings of the BASC-2 subscales, skewness, and kurtosis of the distributions were reviewed. Skewness refers to the tendency of scores clustering on one side of the mean (Jaccard & Becker, 2002). Kurtosis describes the shape of the distribution in terms of height and flatness (Urdan, 2010). The minimum raw scores, maximum raw scores, skewness, and kurtosis statistics were calculated with the SPSS computer statistic program (see Appendix E Table E1 to Table E16). Z-scores were hand calculated for the skewness and kurtosis by dividing these by the standard errors (Kim, 2013) (see Appendix E Table E1 to Table E16).

Most distributions were normally distributed on the BASC-2 subscales. Those demonstrating non-normality as indicated by the skewness statistic, skewness statistic, or z-score included a variety of rating scales and ages from Adaptability, Aggression, Anxiety, Attention Problems, Atypicality, Conduct Problems, Depression, Functional Communication, Hyperactivity, Interpersonal Relations, Learning Problems, Somatization, Study Skills, and Withdrawal (see Table 9).

Table 9
Non-Normal Distributions

Subscale	Rating Scale
Adaptability	TRS-A (F)
Aggression	PRS-A (F), PRS-P (F), TRS-A (B), TRS-C (F), TRS-P (B)
Anxiety	TRS-A (B), TRS-P (B)
Attention Problems	TRS-A (F)
Atypicality	PRS-A (B), PRS-C (B), PRS-P (F), TRS-A (B) TRS-C (F), TRS-P (B)
Conduct Problems	PRS-A (B), TRS-A (B), TRS-C (F)
Depression	TRS-A (B), TRS-C (F), TRS-P (B)
Functional Communication	TRS-A (F)
Hyperactivity	TRS-A (B), TRS-P (B)
Interpersonal Relations	SRP-C (F)
Learning Problems	TRS-A (F)
Somatization	PRS-P (F), TRS-A (B), TRS-C (F), TRS-P (F)
Study Skills	TRS-A (B)
Withdrawal	TRS-A (B), TRS-P (B)

F=Female; M=Male; B=Both

The assumption of normality for independent equivalent *t*-tests was violated by these rating scales and may have affected the results. This might impact the conclusions of similarity drawn from the equivalence tests. Reynolds and Kamphaus (2004) discussed the decision to preserve the non-normal distributions of the BASC-2 rating scales under the belief that problem behaviors in the population maybe distributed in such a manner. This method of standardization is termed a linear transformation in which it is assumed that "the distances between scale points reflect true differences in the population" (p. 128). During the development and standardization of the BASC-2 rating scales, parametric statistics were used extensively 2004). With these considerations in mind the independent *t*-tests of equivalence were conducted on all the BASC-2 rating scales.

Effect sizes. An effect size captures the magnitude of difference or relationship strength between variables (Cooper, 2010). Effect size measures the magnitude of a

treatment or how different two groups are from one another (Salkind, 2008). To compute an effect size the mean difference is divided by a pooled standard deviation (Coladarciet al., 2008). Cohen (1992) developed the operational definitions of small, medium, and large effect sizes. Medium effect sizes were intended to be visible to the naked eye, small were noticeably smaller than medium but not trivial, and large was the same distance from medium as medium was from small (Cohen, 1992). For independent t tests, Cohen proposed effects sizes of .20 for small, .50 for medium, and .80 for large. Effect sizes for this current study are the mean ratings for males minus the ratings for females divided by the pooled standard deviation.

Olejnik and Algina (2000) summarized the criticisms and disadvantages that may be associated with the use of effect sizes. These include how effect sizes actually contribute to a deeper understanding or practical implications of results; the standards of interpretation; the impact of potential research design; and the interpretation of effect sizes in comparison to effect sizes from previous studies (Olejnik & Algina, 2000). Although effect sizes attempt to provide a means by which to examine the meaningfulness of differences it does not provide insight or implications for the findings for practical applications. The magnitude of differences does not indicate what to do with such differences when they are found. As discussed above, Cohen (1992) suggested various guidelines for interpretation of effect sizes. These are often used as the standard for interpretation without consideration of the topic being studied as recommended by Weber and Popova (2012). Also, effect sizes are impacted by the research design and these should be taken into account when calculating the appropriate effect size. For example, different effect size computations should be used depending on the number

(univariate versus multivariate study designs) and type of study variables (Olejnik & Algina, 2000). Lastly, effect sizes are compared across studies; however, reliability, population heterogeneity, levels of variables, and the treatment studies impact the appropriateness of comparing effect sizes of various studies (Olejnik & Algina, 2000).

Pearson product-moment correlation coefficient. A statistic used in this study to examine the relationship between the gender of the student and self, parent, and teacher ratings was Pearson *r*. The Pearson product-moment correlation coefficient (Pearson *r)* is a statistic that indicates the strength and direction of a relationship between two variables, or how much scores on one variable correspond with scores on another variable (Urdan, 2010). An important aspect of this type of statistic is that correlation is not causation (Coladarci, Cobb, Minium, & Clarke, 2008; Urdan, 2010). A type of Pearson correlation coefficient that can be used when one variable is categorical and the other is continuous is a point biserial correlation coefficient (Urdan, 2010). In this study the gender of students was the categorical variable and the ratings were the continuous variables. The point biserial correlation coefficient was calculated for the item analysis on the BASC-2. Guidelines for Pearson *r* include .2 as low, .3-.4 as moderately low, .5-.7 as moderately high, and .8 or higher as strong, and 1 is a perfect correlation.

Coefficient of determination and coefficient of nondetermination. Additional statistics calculated to examine the variability between genders and ratings were the coefficient of determination and nondetermination. Salkind (2008) defined the coefficient of determination as the percentage of variance in one variable that is accounted for by the variance in another variable. In the present study this referred to what percentage gender accounted for the rating on index scores on the BASC-2 rating

scales. The coefficient of nondetermination is the amount, or percentage, of variance that is unexplained (Salkind, 2008) and is conveyed by the overlap between the genders' ratings. The coefficient of determination was calculated by squaring Pearson *r*. The amount of overlap, or unexplained variance, was then determined by subtracting the coefficient of determination from one.

Purpose

The purpose of this study is to examine how the genders are perceived as functioning by student, parent, and teacher ratings on a behavioral, social, and emotional rating scale, specifically the possible similarities between how genders are rated. The statistical analyses conducted are included in the order in which they were performed (see Table 10).

Table 10
Statistical Analyses Performed in Order

1. Group Sample Size
2. Group Means
3. Standard Deviations
4. Pearson's Correlation Coefficient
5. Coefficient of Determination
6. Independent *t*-Tests of Equivalence

Hypotheses:

1. The genders are similarly rated on *the Behavior Assessment System for Children, second edition Parent Rating Scale* (BASC-2 PRS).
2. The genders are similarly rated on *the Behavior Assessment System for Children, second edition Teacher Rating Scale* (BASC-2 TRS).
3. The genders are similarly rated on *the Behavior Assessment System for Children, second edition Self Report of Personality* (BASC-2 SRP).

Chapter 4: Results

The normative data from the BASC-2 ratings scales were analyzed. The mean, standard deviation, Pearson's correlation coefficient, and coefficient of determination were calculated for each subscale on each rating scale. Then independent *t*-tests of equivalence were conducted on each subscale to examine the similarity between the respondents and the gender of the students. These statistics and tests results are provided for each rating scale's subscales.

Table 11
Statistical Analyses Performed in Order

1. Group Sample Size
2. Group Means
3. Standard Deviations
4. Pearson's Correlation Coefficient
5. Coefficient of Determination
6. Independent T-Tests of Equivalence

Adaptive Subscales

The adaptive subscales on the BASC-2 include items that target students' strengths and skills and positive adjustment. These positive attributes and behaviors were combined into the subscales of Adaptability, Activities of Daily Living, Functional Communication, Interpersonal Relations, Leadership, Relations with Parents, Self-Esteem, Self Reliance, Social Skills, and Study Skills.

Preschool Rating Scales: Adaptive Subscales

Parent Rating Scale Preschool (PRS-P). On *the Parent Rating Scale Preschool* (PRS-P) the mean ratings for both genders were less than 2 points different on all adaptive subscales (see Table 12). Although, the correlation between gender and Social Skills on the PRS-P was found to be significant at the 0.01 level, only 2% of the

variability in Social Skills on the PRS-P was accounted for by gender. Correlations on all other adaptive subscales on the PRS-P were not significant.

Table 12
Parent Rating Scale Preschool (PRS-P) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Adaptability	Female	359	15.1	4.3	-.022	.0005
	Male	387	14.9	4.6		
Activities of Daily	Female	359	16.1	4.7	-.119	.0141
Living	Male	387	14.9	5.4		
Functional	Female	359	19.1	6.6	-.079	.0062
Communication	Male	387	18.1	6.8		
Social Skills	Female	359	17.1	5.4	-.139**	.0193
	Male	387	15.5	5.9		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. I also performed a Weber and Popova (2012) independent sample equivalence test to compare ratings of students' genders and the adaptive subscales of the PRS-P (see Table 13). The minimum substantial effect, or delta (Δ), was used based on Cohen's *d* and are conservative guidelines for effect size interpretation at the 0.10, 0.30, and 0.50 levels (Cohen, 1992). At the .10 delta level, or a small effect size, the genders were not rated similarly on Activities of Daily Living, Functional Communication, and Social Skills. However, females and males were rated similarly on all of the adaptive subscale at the .30 and .50 delta levels at the moderate and large effect sizes.

Table 13
Tests of Equivalence Parent Rating Scale Preschool (PRS-P)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Adaptability	.055	.000	.000	.59
Activities of Daily Living	.860	.000	.000	3.28
Functional Communication	.486	.000	.000	2.16
Social Skills	.947	.002	.000	3.82

Teacher Rating Scale Preschool (TRS-P). Much like the PRS-P ratings, the teacher ratings on *the Teacher Rating Scale Preschool* (TRS-P) for female and male students were very similar with mean scores that varied by less than 2 points (see Table 14). Among the subscales, the correlation between gender and Adaptability ratings were found to be significant at the 0.05 level and the correlation between the Social Skills ratings and gender were found to be significant at the 0.01 level. On the Adaptability subscale, 1% of the variation was accounted for by gender and on the Social Skills subscale 1.6% of the variation was accounted for by gender on the TRS-P. No significant correlation was found between gender and the ratings on Functional Communication.

Table 14
Teacher Rating Scale Preschool (TRS-P) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Adaptability	Female	241	13.3	5.1	-.105*	.011
	Male	210	12.3	4.9		
Functional	Female	241	14.9	5.8	-.073	.0053
Communication	Male	210	14.0	6.4		
Social Skills	Female	241	9.8	4.5	-.157**	.025
	Male	210	8.3	4.8		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. Based on the independent t-tests of equivalence the teacher ratings on the adaptive subscales on the TRS-P were similar between genders at the .30 and .50 levels of delta (see Table 15). At the .10 delta level, or small effect size, the genders were not rated similarly on any of the TRS-P Adaptive subscales.

Table 15
Tests of Equivalence Teacher Rating Scale Preschool (TRS-P)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Adaptability	.699	.001	.000	2.23
Functional Communication	.443	.000	.000	1.56
Social Skills	.951	.031	.000	3.37

Child Rating Scales: Adaptive Subscales

Parent Rating Scale Child (PRS-C). On *the Parent Rating Scale Children* (PRS-C) the mean ratings were similar on Adaptability, Activities of Daily Living, and Social Skills. On the Functional Communication and Leadership subscales, the means were more than 2 points different (see Table 16). All adaptive subscales were found to be significantly correlated with gender at the 0.01 level. However, these correlations were not substantively strong and significance is due to a large sample size rather than a true correlational relationship. The coefficients of determination for the PRS-C were calculated and gender accounted for 5.7% of variability on Activities of Daily Living and 5.6% of variability on Social Skills. The coefficients of determinations on the other adaptive subscale ratings on Adaptability, Functional Communication, and Leadership were less than 5%.

Table 16
Parent Rating Scale Child (PRS-C) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Adaptability	Female	501	16.0	4.6	-.160**	.026
	Male	531	14.5	4.8		
Activities of Daily	Female	501	16.1	3.9	-.239**	.057
Living	Male	531	14.2	4.1		
Functional	Female	501	27.1	6.0	-.204**	.042
Communication	Male	531	24.4	6.9		
Leadership	Female	501	14.7	4.7	-.206**	.042
	Male	531	12.7	4.8		
Social Skills	Female	501	14.6	4.7	-.237**	.056
	Male	531	14.7	5.2		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The results of the independent *t*-tests of equivalence were similar at the largest delta 0.50 level for all adaptive subscales on the PRS-C (see Table 17). At the .10 delta level, or small effect size, the genders were not rated similarly on

any of the Adaptive subscales. At the .30 delta level, or moderate effect size, the genders were rated differently on Activities of Daily Living and Social Skills.

Table 17
Tests of Equivalence Parent Rating Scale Child (PRS-C)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Adaptability	.995	.003	.000	5.19
Activities of Daily Living	1.000	.476	.000	7.88
Functional Communication	1.000	.108	.000	6.69
Leadership	1.000	.121	.000	6.76
Social Skills	1.000	.455	.000	7.83

Teacher Rating Scale Child (TRS-C). There was more variation in mean ratings on *the Teacher Rating Scale Child* (TRS-C) adaptive subscales than the preschool rating scales. The Social Skills and Study Skills subscales varied more than the other adaptive subscales on the TRS-C by gender (see Table 18). The correlations were found to be significant at the 0.01 significance level on Adaptability, Functional Communication, Leadership, and Social Skills and at the 0.05 significance level on Study Skills. Similar to the correlations on the other rating scales, these correlations are not strong or substantive and most likely due to the large sample size. The largest coefficient of determination was on the Study Skills subscale at which 4.1% of the variation on subscale was due to gender.

Table 18
Teacher Rating Scale Child (TRS-C) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Adaptability	Female	411	17.3	4.8	-.176**	.031
	Male	378	15.5	5.2		
Functional	Female	411	22.5	5.8	-.139**	.019
Communication	Male	378	20.8	6.4		
Leadership	Female	411	10.3	4.3	-.159**	.025
	Male	378	9.0	4.1		
Social Skills	Female	411	15.1	5.6	-.184**	.034
	Male	378	13.0	5.6		
Study Skills	Female	411	13.7	5.1	-.203*	.041
	Male	378	11.6	5.0		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The genders were rated similarly on the TRS-C on all the adaptive subscales at the .30 and .50 delta levels as calculated by the independent *t*-test of equivalence (see Table 19). The genders were not rated similarly on the Adaptive subscales at the .10 delta level.

Table 19
Tests of Equivalence Teacher Rating Scale Child (TRS-C)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Adaptability	.997	.028	.000	5.02
Functional Communication	.953	.001	.000	3.94
Leadership	.990	.010	.000	4.61
Social Skills	.998	.046	.000	5.25
Study Skills	1.000	.131	.000	5.81

Self Report of Personality Child (SRP-C). Students' mean ratings on *the Self Report of Personality Child* (SRP-C) were very similar and all less than a point different (see Table 20). Correlations between gender and the adaptive subscales on the SRP-C were significant at the 0.05 level for Interpersonal Relations and Relations with Parents and at the 0.01 level for Self Reliance. However, when taking the coefficients of determination into consideration, these correlations are not substantively strong. For

example, only 2.96% of variation in Self Reliance is explained by gender on the SRP-C; 0.9% of variation in Relations with Parents is explained by gender; and 0.7% of variation in Interpersonal Relations is explained by gender. On the Self-Esteem subscale no significant correlation was found.

Table 20
Self Report of Personality Child (SRP-C) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Interpersonal	Female	333	13.9	2.8	-.084*	.0071
Relations	Male	371	13.5	3.0		
Relations with	Female	333	21.1	5.2	-.095*	.0090
Parents	Male	371	20.1	5.2		
Self-Esteem	Female	333	16.3	3.4	.012	.00014
	Male	371	16.4	3.1		
Self Reliance	Female	333	15.9	3.9	-.172**	.0296
	Male	371	14.5	4.0		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The independent t-tests of equivalence conducted on the SRP-C indicated no significant difference at the .30 and .50 delta levels (see Table 21). The genders were rated differently at the smallest effect delta level .10 on the Interpersonal Relations, Relations with Parents, and Self-Reliance.

Table 21
Tests of Equivalence Self Report of Personality Child (SRP-C)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Interpersonal Relations	.540	.000	.000	-2.23
Relations with Parents	.658	.000	.000	-2.54
Self-Esteem	.034	.000	.000	.31
Self-Reliance	.993	.027	.000	-4.62

Adolescent Rating Scales: Adaptive Subscales

Parent Rating Scale Adolescent (PRS-A). On *the Parent Rating Scale Adolescent* (PRS-A) parents' mean ratings on female and male students were most similar on the Adaptability, Activities of Daily Living, and Functional Communication

subscales (see Table 22). The subscales with larger mean differences of almost 2 points as rated by teachers included Leadership and Social Skills. The largest difference between the means of parents' ratings was on the Leadership subscale.

Table 22
Parent Rating Scale Adolescent (PRS-A) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Adaptability	Female	513	15.8	4.6	-.060	.004
	Male	502	14.9	5.0		
Activities of Daily	Female	513	15.5	4.6	-.119**	.0142
Living	Male	502	13.6	4.6		
Functional	Female	513	26.9	6.5	-.079*	.0062
Communication	Male	502	25.2	7.0		
Leadership	Female	513	18.6	6.2	-.196**	.0384
	Male	502	16.0	6.4		
Social Skills	Female	513	16.6	5.0	-.226**	.051
	Male	502	14.2	5.6		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The independent t-tests of equivalence conducted on PRS-A indicated similarity on all of the adaptive subscales on the PRS-A at the delta .50 level (see Table 23). Similarity was found between the genders on all Adaptive subscales, but Social Skills at the .30 delta level. At the .10 delta level the genders were not rated similarly on any of the subscales.

Table 23
Tests of Equivalence Parent Rating Scale Adolescent (PRS-A)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Adaptability	.260	.000	.000	1.91
Activities of Daily Living	1.000	.107	.000	6.62
Functional Communication	.940	.000	.000	4.12
Leadership	1.000	.068	.000	6.37
Social Skills	1.000	.314	.000	7.39

Teacher Rating Scale Adolescent (TRS-A). On *the Teacher Rating Scale Adolescent* (TRS-A) teachers' mean ratings on female and male students were most

similar on the Adaptability and Functional Communication subscales (see Table 24). The subscales with larger mean differences of almost 2 points as rated by teachers included Social Skills and Study Skills. The largest difference between the means of teachers' ratings was on the Study Skills subscale.

The correlations between gender and all the adaptive subscales on the TRS-A were found to be significant at the 0.01 level two-tailed. However, the correlations are not substantively strong. The largest correlation coefficient was on Study Skills for which 7.7% of gender explained the variation in Study Skills on the TRS-A. On the Social Skills subscale 5.7% of the variation in ratings was explained by gender. All other subscales were less than 5%.

Table 24
Teacher Rating Scale Adolescent (TRS-A) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Adaptability	Female	272	17.6	5.0	-.162**	.0262
	Male	272	15.9	5.5		
Functional	Female	272	18.6	4.8	-.165**	.0272
Communication	Male	272	16.9	5.2		
Leadership	Female	272	10.2	4.4	-.180**	.0324
	Male	272	8.5	4.6		
Social Skills	Female	272	15.2	5.7	-.239**	.0571
	Male	272	12.3	6.1		
Study Skills	Female	272	18.6	6.4	-.277**	.0767
	Male	272	14.5	7.7		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. Based on these deltas all of the adaptive subscale ratings were similar for females and males at the large effect size of .50 (see Table 25). At the moderate effect size, .30, the genders were not rated similarly on Social Skills. Again, at the smallest effect size, .10, no subscales were rated similarly between the genders.

Table 25
Tests of Equivalence Teacher Rating Scale Adolescent (TRS-A)

Subscale	Delta .10 *p*, two-tailed	Delta .30 *p*, two-tailed	Delta .50 *p*, two-tailed	*t* value
Adaptability	.974	.027	.000	3.82
Functional Communication	.977	.031	.000	3.89
Leadership	.991	.067	.000	4.26
Social Skills	1.000	.484	.000	5.73
Study Skills	1.000	.822	.000	6.71

Self Report of Personality Adolescent (SRP-A). Students' mean ratings on *the Self Report of Personality Adolescent* (SRP-A) were all less than 2 points different (see Table 26). The largest difference on the mean ratings was on Self-Esteem with a 1.7 rating difference. The correlations between gender and the adaptive subscales of Self-Esteem and Self-Reliance were significant at the 0.01 significance level. Similar to all the other adaptive subscale correlations, these may be significant although they are not substantive. Gender was found to account for less than 3% of the variation of the subscales of Self-Esteem and Self-Reliance on the SRP-A. No significant correlations were found on the Interpersonal Relations and Relations with Parents subscales.

Table 26
Self Report of Personality Adolescent (SRP-A) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Interpersonal	Females	568	15.5	3.3	-.033	.0011
Relations	Male	576	15.3	3.2		
Relations with	Female	568	20.3	6.3	-.041	.0017
Parents	Male	576	19.8	5.9		
Self-Esteem	Female	568	14.8	4.8	.168**	.0282
	Male	576	16.2	3.5		
Self Reliance	Female	568	15.9	3.6	-.173**	.0299
	Male	576	14.6	3.8		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The self ratings on all the adaptive subscales on the SRP-A that were examined with the independent t-tests of equivalence were found to be similar

at the .30 and .50 delta levels (see Table 27). Similarity was found between the genders on Interpersonal Relations and Relations with Parents at the .10 level, but not on Self-Esteem or Self-Reliance.

Table 27
Tests of Equivalence Self Report of Personality Adolescent (SRP-A)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Interpersonal Relations	.055	.000	.000	1.12
Relations with Parents	.094	.000	.000	1.40
Self-Esteem	.999	.005	.000	-5.76
Self Reliance	.999	.008	.000	5.94

Clinical Subscales

The clinical subscales on the BASC-2 include items that target students' problem behaviors. These concerning behaviors were combined into the subscales of Aggression, Anxiety, Attention Problems, Attitude to School, Attitude to Teachers, Atypicality, Conduct Problems, Depression, Hyperactivity, Learning Problems, Locus of Control, Sensation Seeking, Sense of Inadequacy, Social Stress, Somatization, and Withdrawal.

Preschool Rating Scales: Clinical Subscales

Parent Rating Scale Preschool (PRS-P). Parent mean ratings on *the Parent Rating Scale Preschool* (PRS-P) clinical subscales were similar and less than 2 points different on all subscales (see Table 28). The largest mean difference on the PRS-P was on the Atypicality clinical subscale at a 1.6 point difference. The Depression mean ratings were the same for females and males. Correlations were found to be significant at the 0.05 significance level between gender and parent ratings on Aggression; however, less than a percent (0.06%) of variation in parent ratings on the Aggression subscale was accounted for by gender. Attention Problems and Hyperactivity ratings were found to have a significant correlation with gender at the 0.01 significance level. Less than 1% of

variation in both Hyperactivity and Attention Problem parent ratings were accounted for by gender on PRS-P. No significant correlations were found on the other clinical subscales on the PRS-P.

Table 28
Parent Rating Scale Preschool (PRS-P) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Aggression	Female	359	5.0	3.7	.079*	.0006
	Male	387	5.6	4.3		
Anxiety	Female	359	7.5	4.8	-.023	.0005
	Male	387	7.3	5.3		
Attention	Female	359	6.3	3.2	.097**	.0009
Problems	Male	387	7.0	3.5		
Atypicality	Female	359	3.6	3.5	.051	.0026
	Male	387	4.0	3.9		
Depression	Female	359	7.7	3.9	.007	.000049
	Male	387	7.7	4.2		
Hyperactivity	Female	359	9.8	5.2	.098**	.0096
	Male	387	10.9	5.7		
Somatization	Female	359	6.1	3.9	.010	.0001
	Male	387	6.2	4.3		
Withdrawal	Female	359	8.8	4.5	.005	.000025
	Male	387	8.8	4.9		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. Equivalence tests performed on the PRS-P indicated that parent ratings on the PRS-P were similar for both genders at the .30 delta and .50 levels in all clinical subscales (see Table 29). Similarity was found between the genders at the .10 level on the Anxiety, Depression, Somatization, and Withdrawal subscales.

Table 29
Tests of Equivalence Parent Rating Scale Preschool (PRS-P)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Aggression	.485	.000	.000	-2.16
Anxiety	.058	.000	.000	.62
Attention Problems	.674	.000	.000	-2.65
Atypicality	.211	.000	.000	-1.39
Depression	.023	.000	.000	-.19
Hyperactivity	.689	.000	.000	-2.69
Somatization	.028	.000	.000	-.28
Withdrawal	.020	.000	.000	-.13

Teacher Rating Scale Preschool (TRS-P). The mean teacher ratings on *the Teacher Rating Scale Preschool* (TRS-P) clinical subscales were all less than two points different (see Table 30). The largest mean rating difference was on the Hyperactivity subscale at 1.9 points. Attention Problems and gender were found to have a correlation that was significant at the 0.05 level. Correlations between gender and Aggression and Hyperactivity were found to be significant at the 0.01 significant level. However, despite these significant correlations 1% of the variation in Attention Problems was accounted for by gender on the TRS-P and approximately 3% of the variation in Aggression and Hyperactivity ratings were accounted for by gender. No significant correlations were found on any of the other clinical subscales on the TRS-P, including Anxiety, Atypicality, Depression, Somatization, and Withdrawal.

Table 30
Teacher Rating Scale Preschool (TRS-P) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Aggression	Female	241	3.3	4.2	.182**	.033
	Male	210	5.2	5.8		
Anxiety	Female	241	3.8	3.1	.006	.000036
	Male	210	3.8	3.2		
Attention	Female	241	5.8	4.4	.120*	.014
Problems	Male	210	6.8	4.4		
Atypicality	Female	241	2.3	3.4	.039	.0015
	Male	210	2.6	3.4		
Depression	Female	241	3.9	3.5	.076	.0058
	Male	210	4.5	3.9		
Hyperactivity	Female	241	4.7	4.4	.192**	.037
	Male	210	6.6	5.5		
Somatization	Female	241	3.9	3.3	-.077	.0059
	Male	210	3.4	3.0		
Withdrawal	Female	241	4.2	3.2	.037	.0014
	Male	210	4.5	3.7		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. Similar to the results from the independent t-tests of equivalence on the PRS-P, the results for the TRS-P indicate similar ratings for both genders on the TRS-P at the .30 and .50 delta levels (see Table 31). Similar ratings were found only on the Anxiety at the .10 delta small effect size.

Table 31
Tests of Equivalence Teacher Rating Scale Preschool (TRS-P)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Adaptability	.699	.001	.000	2.23
Aggression	.986	.094	.000	-3.92
Anxiety	.058	.000	.000	-.14
Attention Problems	.801	.004	.000	-2.55
Atypicality	.189	.000	.000	-.82
Depression	.460	.000	.000	-1.60
Functional Communication	.443	.000	.000	1.56
Hyperactivity	.992	.134	.000	-4.14
Social Skills	.951	.031	.000	3.37
Somatization	.469	.000	.000	1.63
Withdrawal	.181	.000	.000	-.79

Child Rating Scales: Clinical Subscales

Parent Rating Scale Child (PRS-C). On *the Parent Rating Scale Child* (PRS-C) parents' mean ratings on female and male students were most similar on the Aggression, Anxiety, Atypicality, Depression, Somatization, and Withdrawal subscales (see Table 32). The subscales with larger mean differences of 2 points as rated by parents included Attention Problems and Hyperactivity, these were also the largest difference in ratings. The correlations between gender and parent ratings on Aggression, Attention Problems, Atypicality, Conduct Problems, Depression, and Hyperactivity were found to be significant at the 0.01 significance level. Withdrawal parent ratings and gender correlation was found to be significant at the 0.05 level. The strongest coefficient of determination was found for Attention Problems for which gender accounted for 6.2% of variation on the PRS-C. No significant correlations were found on the PRS-C between gender and Anxiety and Somatization.

Table 32
Parent Rating Scale Child (PRS-C) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Aggression	Female	501	4.4	3.7	.171**	.029
	Male	531	5.9	4.6		
Anxiety	Female	501	12.5	6.1	-.026	.000676
	Male	531	12.2	6.3		
Attention	Female	501	5.0	3.8	.249**	.062
Problems	Male	531	7.0	4.0		
Atypicality	Female	501	3.0	3.8	.144**	.021
	Male	531	4.3	4.9		
Conduct Problems	Female	501	4.6	3.4	.127**	.016
	Male	531	5.6	3.9		
Depression	Female	501	5.9	5.1	.094**	.009
	Male	531	7.0	6.3		
Hyperactivity	Female	501	6.8	4.7	.195**	.038
	Male	531	8.8	5.6		
Somatization	Female	501	5.3	4.0	-.057	.003
	Male	531	4.9	4.1		
Withdrawal	Female	501	6.6	4.7	.078*	.006
	Male	531	7.3	5.5		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. When parent ratings of female and male children were examined with the independent t-tests of equivalence it was found that all parent ratings on the PRS-C were similar at the most conservative minimum substantive effect of delta .50 (see Table 33). Similar ratings were found on most ratings at the .30 delta level, with an exception on the Attention Problems subscale. Similarity was found only on the Anxiety subscale at the .10 level.

Table 33
Tests of Equivalence Parent Rating Scale Child (PRS-C)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Aggression	.999	.009	.000	-5.57
Anxiety	.041	.000	.000	.84
Attention Problems	1.000	.622	.000	-8.26
Atypicality	.982	.001	.000	-4.68
Conduct Problems	.938	.000	.000	-4.12
Depression	.669	.000	.000	-3.02
Hyperactivity	1.000	.062	.000	-6.39
Somatization	.228	.000	.000	1.83
Withdrawal	.473	.000	.000	-2.51

Teacher Rating Scale Child (TRS-C). On *the Teacher Rating Scale Child* (TRS-Child) teachers' mean ratings on female and male students were most similar on the Aggression, Anxiety, Atypicality, Conduct Problems, Depression, Learning Problems, Somatization, and Withdrawal subscales (see Table 34). Some of the largest differences in mean teacher ratings were found on the Attention Problems clinical subscale (3.2) and the Hyperactivity clinical subscale (4.1).

Correlations between gender and Aggression, Attention Problems, Atypicality, Conduct Problems, Hyperactivity, Learning Problems, and Withdrawal were significant at the 0.01 significance level. The correlation between the Depression subscale ratings and gender were found to be significant at the 0.05 significance level. However, most of these correlations are weak and not substantively strong. Teachers' ratings on Hyperactivity and gender indicated a small relationship. Most of the coefficients of determination were low. Gender was found to account for 8.4% of variation in Attention Problem ratings and 9.5% of Hyperactivity ratings on the TRS-C. Similar to the ratings on the PRS-C, there were no significant correlations on Anxiety or Somatization on the TRS-C.

Table 34
Teacher Rating Scale Child (TRS-C) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Aggression	Female	411	2.0	3.4	.226**	.051
	Male	378	3.9	4.9		
Anxiety	Female	411	4.0	3.3	-.053	.003
	Male	378	3.7	3.4		
Attention	Female	411	5.6	5.1	.289**	.084
Problems	Male	378	8.8	5.5		
Atypicality	Female	411	1.6	2.6	.138**	.019
	Male	378	2.4	3.4		
Conduct Problems	Female	411	2.4	3.5	.175**	.031
	Male	378	3.7	3.9		
Depression	Female	411	2.4	3.4	.086*	.007
	Male	378	3.2	3.8		
Hyperactivity	Female	411	4.8	5.4	.309**	.095
	Male	378	8.9	7.2		
Learning Problems	Female	411	3.7	4.3	.160**	.026
	Male	378	5.2	5.0		
Somatization	Female	411	2.5	3.5	-.060	.0036
	Male	378	2.1	2.9		
Withdrawal	Female	411	3.3	3.5	.112**	.013
	Male	378	4.2	4.2		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The results of the tests of equivalence on the TRS-C indicated that all of the teacher ratings on the genders on the clinical subscales were similar at the .30 and .50 delta levels (see Table 35). The genders were rated differently at the .30 delta level on the Aggression, Attention Problems, and Hyperactivity.

Table 35
Tests of Equivalence Teacher Rating Scale Child (TRS-C)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Aggression	1.000	.335	.000	-6.52
Anxiety	.221	.000	.000	1.48
Attention Problems	1.000	.930	.000	-8.46
Atypicality	.950	.001	.000	-3.90
Conduct Problems	.997	.026	.000	-4.99
Depression	.564	.000	.000	-2.42
Hyperactivity	1.000	.982	.001	-9.10
Learning Problems	.988	.008	.000	-4.53
Somatization	.283	.000	.000	1.68
Withdrawal	.814	.000	.000	-3.15

Child Rating Scales: Clinical Subscales

Self Report of Personality Child (SRP-C). On *the Self Report of Personality Child* (SRP-C) students' mean ratings were similar on most of the clinical subscales (see Table 36). A 2.4 mean difference was found on the Attention Problems subscale for male and female students self ratings. Many clinical subscale and gender correlations were found to be significant at the 0.05 and 0.01 levels, but no correlations were substantively strong. The largest coefficient of determination was found for the Attention Problems subscale on which gender accounted for 5.6% of the variation in ratings on the SRP-C. All other subscale coefficients of determination were less than 5%. The correlation between gender and the ratings on Anxiety subscale was not significant.

Table 36
Self Report of Personality Child (SRP-C) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Anxiety	Female	333	11.5	7.1	.013	.000169
	Male	371	11.7	6.2		
Attention	Female	333	6.5	4.7	.236**	.0557
Problems	Male	371	8.9	5.0		
Attitude to School	Female	333	4.4	4.0	.209**	.0437
	Male	371	6.3	4.7		
Attitude to	Female	333	2.8	3.3	.124**	.0154
Teachers	Male	371	3.7	3.5		
Atypicality	Female	333	5.1	5.3	.117**	.0137
	Male	371	6.4	5.1		
Depression	Female	333	4.4	5.3	.118**	.0139
	Male	371	5.6	5.1		
Hyperactivity	Female	333	6.7	4.7	.150**	.0225
	Male	371	8.1	4.6		
Locus of Control	Female	333	5.1	4.3	.103**	.0106
	Male	371	5.9	4.1		
Sense of	Female	333	5.1	3.9	.105**	.0110
Inadequacy	Male	371	5.8	3.5		
Social Stress	Female	333	5.1	4.3	.090*	.0081
	Male	371	5.6	4.1		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The independent t-tests of equivalence indicated similarity on all the clinical subscales on the SRP-C ratings at the delta .50 level (see Table 37). Most of the students' ratings were similar at the moderate .30 level, with the exception of Attention Problems. At the small effect size delta level of .10 similarity was only found on Anxiety.

Table 37
Tests of Equivalence Self Report of Personality Child (SRP-C)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Anxiety	.038	.000	.000	.35
Attention Problems	1.000	.451	.000	6.44
Attitude to School	1.000	.186	.000	5.66
Attitude to Teachers	.882	.001	.000	3.32
Atypicality	.839	.000	.000	3.13
Depression	.845	.000	.000	3.15
Hyperactivity	.969	.006	.000	4.01
Locus of Control	.727	.000	.000	2.73
Sense of Inadequacy	.748	.000	.000	2.80
Social Stress	.601	.000	.000	2.39

Adolescent Rating Scales: Clinical Subscales

Parent Rating Scale Adolescent (PRS-A). On *the Parent Rating Scale Adolescent* (PRS-A) parents' mean ratings on female and male students were similar on most of the clinical subscales (see Table 38). Parent mean ratings varied the most on the Anxiety clinical subscale with a 2.3 difference. Correlations with gender and Anxiety, Attention Problems, Depression, and Somatization were found to be significant at the 0.01 significance level. However, none of these correlations were substantively strong. Coefficients of determination were low with the highest being 3.6% of variation of Anxiety and 3.5% of variation of Attention Problems accounted for by gender. All other PRS-A clinical subscale ratings correlations with gender were not significant.

Table 38
Parent Rating Scale Adolescent (PRS-A) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Aggression	Female	513	4.9	5.8	-.017	.000289
	Male	502	4.7	4.8		
Anxiety	Female	513	11.8	6.1	-.190**	.036
	Male	502	9.5	5.5		
Attention	Female	513	5.1	3.9	.186**	.035
Problems	Male	502	6.6	3.9		
Atypicality	Female	513	4.2	6.0	-.056	.003
	Male	502	3.6	4.4		
Conduct Problems	Female	513	6.0	7.7	-.001	.000001
	Male	502	6.0	5.6		
Depression	Female	513	7.8	7.6	-.131**	.017
	Male	502	6.0	5.9		
Hyperactivity	Female	513	5.0	4.8	.039	.002
	Male	502	5.3	4.4		
Somatization	Female	513	6.4	6.2	-.151**	.023
	Male	502	4.8	4.6		
Withdrawal	Female	513	5.8	4.4	.027	.000729
	Male	502	6.0	4.7		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. Results from the independent t-tests of equivalence indicated that parents' ratings were similar on all the clinical subscales on the PRS-A at the .30 and .50 delta levels (see Table 39). Similar ratings were found between genders on Aggression, Conduct Problems, Hyperactivity, and Withdrawal at the .10 delta level.

Table 39
Tests of Equivalence Parent Rating Scale Adolescent (PRS-A)

Subscale	Delta .10 *p*, two-tailed	Delta .30 *p*, two-tailed	Delta .50 *p*, two-tailed	*t* value
Aggression	.023	.000	.000	.56
Anxiety	1.00	.043	.000	6.15
Attention Problems	1.000	.033	.000	-6.02
Atypicality	.221	.000	.000	1.79
Conduct Problems	.006	.000	.000	.04
Depression	.949	.000	.000	4.20
Hyperactivity	.095	.000	.000	-1.25
Somatization	.989	.001	.000	4.85
Withdrawal	.044	.000	.000	-.85

Teacher Rating Scale Adolescent (TRS-A). On *the Teacher Rating Scale Adolescent* (TRS-A) teachers' mean ratings on female and male students were most similar on the Anxiety, Atypicality, Depression, Somatization, and Withdrawal subscales. The subscales with larger mean differences of almost 2 points as rated by teachers included Aggression, Attention Problems, Conduct Problems, Hyperactivity, and Learning Problems (see Table 40). Correlations were substantively weak, although, significance was found at the 0.01 level on Aggression, Attention Problems, Atypicality, Conduct Problems, Depression, Hyperactivity, and Withdrawal. The largest coefficient of determination was on Attention Problems for which 8.5% of gender explained the variation in Attention Problems on the TRS-A. No significant correlations were found between gender and Anxiety and Somatization ratings on the TRS-A clinical subscales.

Table 40
Teacher Rating Scale Adolescent (TRS-A) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Aggression	Female	272	1.8	3.1	.229**	.0524
	Male	272	3.7	4.9		
Anxiety	Female	272	3.7	3.4	-.026	.000068
	Male	272	3.6	3.6		
Attention	Female	272	4.3	4.3	.293**	.0858
Problems	Male	272	7.3	5.3		
Atypicality	Female	272	1.2	2.2	.203**	.0412
	Male	272	2.4	3.4		
Conduct Problems	Female	272	2.0	3.5	.233**	.0543
	Male	272	4.0	5.1		
Depression	Female	272	2.3	3.2	.146**	.0213
	Male	272	3.4	4.1		
Hyperactivity	Female	272	3.4	4.4	.250**	.0625
	Male	272	6.3	6.6		
Learning Problems	Female	272	3.0	3.9	.215**	.0462
	Male	272	4.9	4.7		
Somatization	Female	272	1.7	2.7	.024	.000058
	Male	272	1.8	2.8		
Withdrawal	Female	272	3.6	3.5	.144**	.0207
	Male	272	4.8	4.3		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The results from the independent t-tests of equivalence indicated all of the subscale ratings were similar for females and males at the large effect size delta level of .50 (see Table 41). Similar ratings were found at the .10 delta level on Somatization and at the .30 delta level on Anxiety, Atypicality, Depression, Learning Problems, Somatization, and Withdrawal.

Table 41
Tests of Equivalence Teacher Rating Scale Adolescent (TRS-A)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Aggression	1.000	.387	.000	-5.48
Anxiety	.101	.000	.000	.59
Attention Problems	1.000	.907	.001	-7.12
Atypicality	.998	.173	.000	-4.82
Conduct Problems	1.000	.425	.000	-5.58
Depression	.939	.010	.000	-3.43
Hyperactivity	1.000	.593	.000	-6.01
Learning Problems	.999	.264	.000	-5.13
Somatization	.095	.000	.000	-.56
Withdrawal	.935	.009	.000	-3.39

Self Report of Personality Adolescent (SRP-A). On *the Self Report of Personality Adolescent* (SRP-A) students' mean ratings were similar on most of the clinical subscales (see Table 42). Their mean ratings varied more than 2 points on the Anxiety and Sensation Seeking subscales. Correlations between the subscale ratings and gender were weak, although most were indicated as being significant at the 0.01 and 0.05 levels due to a large sample size. The largest coefficients of determination were found to be on the Anxiety and Sensation Seeking subscales. It was calculated that 4.6% of variation on the Anxiety and 5.3% of variation on the Sensation Seeking subscales was accounted for by gender. No significant correlations were found between gender and the SRP-A on Atypicality, Locus of Control, and Sense of Inadequacy.

Table 42
Self Report of Personality Adolescent (SRP-A) Descriptive Statistics

Subscale	Gender	n	*M*	SD	*r*	r^2
Anxiety	Female	568	13.2	7.5	-.214**	.0458
	Male	576	10.2	6.2		
Attention	Female	567	6.8	4.7	.112**	.0125
Problems	Male	576	7.9	4.7		
Attitude to School	Female	568	5.9	3.95	.061*	.0037
	Male	576	6.4	4.1		
Attitude to	Female	568	4.95	4.2	.118**	.0139
Teachers	Male	576	5.95	4.1		
Atypicality	Female	568	3.2	4.2	.046	.0021
	Male	576	3.8	4.1		
Depression	Female	568	5.2	5.8	-.087**	.0076
	Male	576	4.3	4.8		
Hyperactivity	Female	568	6.1	3.7	.059*	.0035
	Male	576	6.6	4.1		
Locus of Control	Female	568	5.8	4.8	-.002	.000004
	Male	576	5.7	4.3		
Sensation Seeking	Female	568	9.6	4.5	.230**	.0529
	Male	576	11.7	4.4		
Sense of	Female	568	7.1	4.9	-.028	.0008
Inadequacy	Male	576	6.8	4.3		
Social Stress	Female	568	6.8	5.3	-.062*	.0038
	Male	576	6.2	4.7		
Somatization	Female	568	3.3	3.6	-.143**	.0204
	Male	576	2.4	2.9		

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Tests of equivalence. The results from the independent t-tests of equivalence indicate the students ratings were similar based on gender across the clinical subscales at the delta .50 level and most subscales at the .30 delta level with exception of Sensation Seeking (see Table 43). Similarity was found across ratings at the .10 delta level on Locus of Control and Sense of Inadequacy.

Table 43
Tests of Equivalence Self Report of Personality Adolescent (SRP-A)

	Delta .10	Delta .30	Delta .50	
Subscale	*p*, two-tailed	*p*, two-tailed	*p*, two-tailed	*t* value
Anxiety	1.000	.174	.000	7.42
Attention Problems	.863	.000	.000	-3.81
Attitude to School	.262	.000	.000	-2.08
Attitude to Teachers	.904	.000	.000	-4.02
Atypicality	.126	.000	.000	-1.57
Depression	.590	.000	.000	2.94
Hyperactivity	.234	.000	.000	-1.99
Locus of Control	.004	.000	.000	.05
Sensation Seeking	1.000	.351	.000	-7.98
Sense of Inadequacy	.038	.000	.000	.94
Social Stress	.264	.000	.000	2.08
Somatization	.985	.000	.000	4.90

Summary

Correlations and coefficients of determination were weak and did not demonstrate substantive strength between the BASC rating scales subscales and gender. Similarities were found at the small effect size delta level of .10 on 18 subscales (see Table 44, Table 45, and Table 46).

Table 44
Tests of Equivalence at the Delta .10 Level on the BASC-2 Similar Ratings

Adaptive	Clinical
Adaptability (PRS-P)	Anxiety (PRS-P, TRS-P, PRS-C, SRP-C)
Self-Esteem (SRP-C)	Depression (PRS-P)
Interpersonal Relations (SRP-A)	Somatization (PRS-P, TRS-A)
Relations with Parents (SRP-A)	Withdrawal (PRS-P, PRS-A)
	Aggression (PRS-A)
	Conduct Problems (PRS-A)
	Hyperactivity (PRS-A)
	Locus of Control (SRP-A)
	Sense of Inadequacy (SRP-A)

Table 45
Tests of Equivalence Similar at delta .10 Adaptive Subscales

		Delta .10	
Subscale	Rating Scale	*p*, two-tailed	*t* value
Adaptability	PRS-P	.055	.59
Self-Esteem	SRP-C	.034	.31
Interpersonal Relations	SRP-A	.055	1.12
Relations with Parents	SRP-A	.094	1.40

Table 46
Tests of Equivalence Similar at delta .10 Clinical Subscales

		Delta .10	
Subscale	Rating Scale	*p*, two-tailed	*t* value
Anxiety	PRS-P	.058	.62
Depression	PRS-P	.023	-.19
Somatization	PRS-P	.028	-.28
Withdrawal	PRS-P	.020	-.13
Anxiety	TRS-P	.058	-.14
Anxiety	PRS-C	.041	.84
Anxiety	SRP-C	.038	.35
Aggression	PRS-A	.023	.56
Conduct Problems	PRS-A	.006	.04
Hyperactivity	PRS-A	.095	-1.25
Withdrawal	PRS-A	.044	-.85
Somatization	TRS-A	.095	-.56
Locus of Control	SRP-A	.004	.05
Sense of Inadequacy	SRP-A	.038	.94

Differences were found between genders at the moderate effect size delta level of .30 on the adaptive subscales of Activities of Daily Living and Social Skills on the PRS-P; Social Skills on the PRS-A; and Social Skills and Study Skills on the TRS-A (see Table 47 and Table 48).

Table 47
Tests of Equivalence at the Delta .30 Level on the BASC-2 Non-Similar Ratings

Adaptive	Clinical
Activities of Daily Living (PRS-C)	Attention Problems (PRS-C, TRS-C, SRP-C, TRS-A)
Social Skills (PRS-C, PRS-A, TRS-A)	Aggression (TRS-C, TRS-A)
Study Skills (TRS-A)	Hyperactivity (TRS-C, TRS-A)
	Conduct Problems (TRS-A)
	Sensation Seeking (SRP-A)

Table 48
Tests of Equivalence Not Similar at delta .30 Adaptive Subscales

		Delta .30	
Subscale	Rating Scale	P, two-tailed	t value
Activities of Daily Living	PRS-C	.476	7.88
Social Skills	PRS-C	.455	7.83
Social Skills	PRS-A	.314	7.39
Social Skills	TRS-A	.484	5.73
Study Skills	TRS-A	.822	6.71

On the clinical subscales genders were not rated similarly at the moderate effect size delta level of .30 on Attention Problems on the PRS-C; Aggression, Attention Problems, and Hyperactivity on the TRS-C; Attention Problems on the SRP-C; Aggression, Attention Problems, Conduct Problems, and Hyperactivity on the TRS-A; and Sensation Seeking on SRP-A (see Table 49).

Table 49
Tests of Equivalence Not Similar at delta .30 Adaptive Subscales

		Delta .30	
Subscale	Rating Scale	P, two-tailed	t value
Activities of Daily Living	PRS-C	.476	7.88
Social Skills	PRS-C	.455	7.83
Social Skills	PRS-A	.314	7.39
Social Skills	TRS-A	.484	5.73
Study Skills	TRS-A	.822	6.71

Overall, based on the independent t-test equivalence testing, the genders were rated similarly on all of the adaptive and clinical subscales at the large effect size delta level of .50. More similarities were found in ratings between the genders on all subscales than differences.

Correlations between gender and subscales increased as the age of the student increased. Similarly, the tests of equivalence weakened as the age of the students increased; more differences in gender were rated as the students were older. More similarities were also rated as the students age increased.

Most adaptive subscales were rated similarly for female and male students by the students, parents, and teachers and a total of 30 subscales out of 35, or a little less than 86% of the adaptive scales, were rated similarly at the .30 and .50 delta levels. The genders were rated similarly on the clinical subscales with 61 out of 76 subscales rated similarly at the .30 delta level. Thus, approximately 80% of the clinical subscales on the BASC-2 were rated similarly at the moderate effect delta levels. Out of the 111 subscales on the BASC-2 rating scales female and male students were rated similarly on all subscales at the.50 delta level strong effect, or 100% of the BASC-2 subscales, and 15 subscales at the .30 delta level moderate effect, or 9% of the BASC-2 subscales. To summarize, at the small effect size of .10 18 out of 111 subscales or 16% equivalence

was found, at the moderate effect size of .30 96 out of 111 subscales or 86% equivalence was found, and at the large effect size .50 111 out of 111 subscales or 100% equivalence was found.

Chapter 5: Discussion

The purpose of this study was to examine how students are perceived as functioning socially, emotionally, and behaviorally by themselves and their parents and teachers on a multi-faceted social, emotional, and behavioral rating scale. The emphasis throughout the study was on examining the similarities between female and male students' ratings. Raw data from the development of the BASC-2 SRP, PRS, and TRS were analyzed for similarity through tests of equivalence. Independent tests of equivalence were performed to examine the similarity between the genders and self, parent, and teacher ratings on the SRP, PRS, and TRS. Tests of equivalence have been used to examine similarity between groups or distributions in performance tests and meta-analyses (Ball et al., 2013; Weber & Popova, 2012). The independent variable of the study was students' gender and the dependent variables were the raw scores on the various scales of the BASC-2.

In this chapter I review and discuss the study results practical, social, and research implications and the study strengths and limitations. Most of the research hypotheses were supported, with the exception of two subscales on the TRS-C and TRS-A. On these subscales small differences were found between female and male students in teachers' ratings. These findings have implications for practice in schools, in our society, and for future research in social, emotional, and behavioral issues.

Hypothesis 1

My first hypothesis was that the genders would be similarly rated on *the Behavior Assessment System for Children, second edition Parent Rating Scale* (BASC-2 PRS). Parents rated the genders similarly on all subscales and at all age levels at the strong

effect size delta level .50. Similarities were found at the smallest effect size level delta .10 on Adaptability (PRS-P), Aggression (PRS-A), Anxiety (PRS-P, PRS-C), Conduct Problems (PRS-A), Depression (PRS-P), Hyperactivity (PRS-A), Somatization (PRS-P), and Withdrawal (PRS-P, PRS-A). At the moderate effect size delta level .30 differences, were found between the genders on Activities of Daily Living (PRS-C), Attention Problems (PRS-C), and Social Skills (PRS-P, PRS-A). Very weak or no relationships were found between students' gender and the subscale ratings. Very little variation in the ratings for students was explained by gender.

Hypothesis 2

The second hypothesis tested was that the genders would be similarly rated on *the Behavior Assessment System for Children, second edition Teacher Rating Scale* (BASC-2 TRS). All subscales were rated similarly at the strong effect 0.50 delta level. This hypothesis was found to be true at the delta .10 level for Anxiety (TRS-P) and Somatization (TRS-A). At the moderate effect size delta level of .30 Aggression (TRS-C, TRS-A), Attention Problems (TRS-C, TRS-A), Hyperactivity (TRS-C, TRS-A), Social Skills (TRS-A), and Study Skills (TRS-A) were rated differently by gender, although, 90% or more of the variation in all these subscales of teaching ratings was not accounted for by gender. Weak relationships were found in teachers' ratings between gender and these subscales with 9% or less of the variation in the ratings explained by gender. Teachers were the respondents who rated the genders most differently across the subscales.

Hyde (2014) cautioned that researchers should be mindful of situational variables along with identifying contexts that affect the appearance or disappearance of gender

differences. For example, situational variables found in the classroom and school campus environments influence students' behavior in schools. With Hyde in mind, the expectations for students in the classroom settings in elementary and at the secondary levels should be examined to understand how situational variables and contexts impact behavioral expectations for female and male students. These expectations may be different than those in other contexts, which might explain the parent and self-ratings that were similar on these subscales and the teachers' ratings that varied by home and school environments.

Hypothesis 3

The final hypothesis tested was that the genders are similarly rated on *the Behavior Assessment System for Children, second edition Self Report of Personality* (BASC-2 SRP). Female and male students rated themselves similarly on all subscales on the BASC-2 at the large effect size delta level .50. At the moderate effect size delta level .30, different ratings were found between the genders on Attention Problems (SRP-C) and Sensation Seeking (SRP-A). Similar ratings were found at the small effect size delta level .10 for students on Self-Esteem (SRP-C), Anxiety (SRP-C), Locus of Control (SRP-A), and Sense of Inadequacy (SRP-A). Similar to parent ratings, very weak to no relation was found between subscale ratings and gender and very little variation in ratings were explained by variation in gender.

Implications

The current findings indicate students are perceived as functioning socially, emotionally, and behaviorally very similarly by themselves, their parents, and their teachers in most areas. Teachers presented that most difference in students' functioning

as compared to students and parents. This has important implications in identifying, addressing, and supporting students' needs. These findings indicate that although many gender stereotypes and the expectations of gender differences exist, those people closest and most present in students' lives do not perceive them as behaving significantly different from one another based on their gender. Additionally these perceived differences are reported less than would be expected based on the emphasis placed on gender differences in student behavior and abilities. For example, female and male students were rated similarly at the moderate and strong effect sizes on Anxiety and Depression, and Self-Esteem at all age levels by students, parents, and teachers.

The present findings have implications for students in the school settings, the practice of school psychologists, and future research directions. The similar ratings of female and male students in schools by the students themselves and teachers provide insight as to how students are viewed as functioning behaviorally, emotionally, and socially. School psychologists receive referrals and conduct assessments of students' behavioral, social, and emotional functioning. Understanding the gender based expectations of mental health disorders and the perception of gender similarities on such behavioral rating scales as the BASC-2 presents important implications in the referral, assessment, and intervention process for school psychologists. Lastly, the use of tests of equivalence to examine similarity between gender groups offers another way to conceptualize and approach research questions, thinking in terms of similarity rather than group differences.

Social Implications. The similarity between the ratings on this broad band measure of social, emotional, and behavioral functioning demonstrates that people

perceive students as displaying similar behaviors despite gender. Socially, gender differences are emphasized and expected through social institutions, including schools (Kimmel, 2011). Yet, this study serves as an example that although gender differences are highlighted in our society they are not necessarily being perceived in students' behavior in the home, school, and community settings. Furthermore, these findings suggest additional review of gender prevalence and symptomology of mental health disorders. Could there be an over identification and under identification of students' needs based on gender stereotyped expectations of behavior?

Internalizing and externalizing mental health disorders. Gender differences in prevalence and symptoms of mental health disorders are emphasized in the DSM-5. Females are reported as being diagnosed more with internalizing type disorders, such as Major Depressive Disorder, Specific Phobias, Social Phobia, and Generalized Anxiety Disorder. However, in the current study gender differences were not found in student, parent, or teacher ratings on internalizing behaviors or disorders. Externalizing mental health disorders reviewed in the DSM-5 included Attention Deficit Hyperactivity Disorder (ADHD), Conduct Disorder (CD), and Oppositional Defiance Disorder (ODD). ADHD and CD are reported as being more common in males than females and ODD is more common in males before puberty, after puberty the prevalence levels for males and females even out (American Psychiatric Association, 2013). Moderate gender differences were found for externalizing type behaviors as rated by mostly teachers and all at the child or adolescent levels. Understanding that internalizing disorders are those not usually associated with behaviors and externalizing disorders are associated with

many behaviors is important to addressing the under and over-identification of students for support services.

Over and under identification of students in need. Male students have long been at-risk for being identified as eligible for special education programs (Coutinho & Oswald, 2005; Sullivan & Bal, 2013). Additionally, female students have been found to be underrepresented in special education programs (Wehmeyer & Schwartz, 2001). Young et al. (2010) found that teachers reported males three times more likely than female students as displaying at risk internalizing and externalizing difficulties. Specifically, male middle school students were five times more likely than female students to be at risk for externalizing behaviors and twice as likely for internalizing at risk behaviors.

Referral concerns have also long existed in the field of school psychology. Such concerns date back into the early emergence of our field of school psychology. Emery (1973) discussed the relationship between special education referrals for hyperactive and aggressive behavior. It was proposed that because girls were socialized to display aggression indirectly and less likely to display disruptive behavior in the classroom, they were more likely to be overlooked in referrals for learning difficulties. On the other hand, boys have been socialized to be aggressive but are penalized for such disruptive behavior in the school setting and are more likely to be referred for hyperactivity. As proposed by Engin, Leppaluoto, and Fodor (1973) consciousness raising steps can be taken on individual and system levels. Conscious raising involves education and informing school staff and families of gender similarities, gender stereotypes, and gender bias that impact students. Highlighting the similarity that genders share and reflecting on

gendered expectations for behavior with students, school staff, and families begins a conversation and lays the groundwork for change in recognizing and supporting students' need despite gender.

Policy and gender similarities. Evidence that students are viewed as functioning similarly across genders has implications for policy. Over-identification of male students receiving special education services or at risk for needing such support indicates a need to document and report gender counts of students to the federal government under IDEIA just as race and ethnicity counts are reported. These statistics are needed nation-wide to understand the gender make up of students receiving special education services to address over and under representation of students' access to services.

Implications for schools. Gender bias, gender stereotypes, and gender expectations are present in our society and school system (Rousso & Wehmeyer, 2001). Gender differences are acknowledged, anticipated, and emphasized at the macro and micro levels of our education system (Hanson & Smith, 2001). For example, Lawrence Summers, a past president of Harvard University spoke at a National Bureau of Economic Research luncheon and presented the hypothesis that innate differences between the genders is the cause for an underrepresentation of female scientists at prestigious universities in addition to aptitude and the ability to dedicate 80 hours weekly to research (Hemel, 2005). His thoughts angered many and were cited as being taken out of context (D.M., 2009; Hemel, 2005). According to D.M. (2005), Summers referred to the perspective that the standard deviation and variability in men's innate mathematical abilities varied more than women's abilities; thus, men's abilities are lower than and exceed women's abilities in math. This hypothesis is very similar to the theories posited

in the early twentieth century by Cattell and Thorndike. Thorndike (1908) proposed that men's abilities had a larger range than female's abilities, and Cattell (1903) observed men held more prominent roles and made more accomplishments in science, technology, and politics. In 100 years several of the cited observations of gender differences in educational and society gender gaps continue to be discussed.

The emphasis on gender differences has been used as the foundational argument for different educational programs to address such proposed differences. One example of such programs is single sex classrooms. Single-sex classrooms, or classes with only female or male students, have been seen as and used to address social and gender inequities and academic achievement in schools (Arms, 2007; Hubbard & Datnow, 2005; Spielhagen, 2011). For example, research has shown that educational experiences vary by gender, race, and ethnicity leading to an increase in interest in single-sex education for various ethnic and racial groups (Hubbard & Datnow, 2005). Furthermore, in 2004 the Bush administration lessened the restrictions of Title IX to allow single-sex classes be offered as options rather than be deemed as a form of sex segregation (Bailey, 2013; Hubbard & Datnow, 2005). Since this, the number of single-sex classrooms has increased from 2007 to 2010 with 1,000 out of 98,000 public kindergarten through 12th grade schools having single-sex academic classrooms (Bailey, 2013). It is most commonly argued that females and males learn differently and would thus benefit from being taught separately to address and accommodate these differences (Spielhagen, 2011; the State Education Resource Center, 2013) and that separating the genders will address distractions and social pressures commonly found in coeducational classes (the State Education Resource Center, 2013). However, in a meta-analysis of 184 studies from 21

countries including 1.6 million K-12 students modest benefits were found for single-sex schooling in mathematics performance outcomes in uncontrolled studies and in controlled studies no advantages were found for single-sex education (Pahlke, Hyde, & Allison, 2014). Thus, the assumption that gender differences exist and should be addressed through single sex classrooms has shown few advantages. The approach of focusing on gender differences to addressing gender equity in education has not been effective; then how should gender equity or students' access to services be framed in terms of gender?

In this study more similarity existed between the genders than difference. Recognizing the similarities between genders in students' functioning is needed in our society. Assuming difference in behavior and needs prevents identifying risks in students' emotional, behavioral, and social functioning. This impacts the referrals for support, the assessment process, and ultimately access to treatment. Most importantly, there is a need for the acknowledgement that those in students' lives view the genders as demonstrating mostly similar behavior and skills despite gender.

Assessing social, emotional, and behavioral functioning. As posed in Chapter 2, should gender differences be taken into consideration when assessing students? Based on the findings of this study, the answer seems to be no. Because social, emotional, and behavioral rating scales are a large component of special education assessments it is necessary to understand how gender is related to the ratings of students, parents, family members, and school staff. Presently, students, parents, and teachers rate students similarly based on gender on adaptive skills and problem behaviors. Reynold and Kamphaus (2004) cautioned that although separate sex norm groups were provided for scoring, they should not be used in all interpretative cases. Yet, no guidelines or

suggestions are provided in the BASC-2 manual as to when it would be appropriate to use the separate sex norm groups. Statistically significant sex differences found during the development of the BASC-2 reached half of a standard deviation. These subscale differences were found on the PRS and TRS (Adaptability, Social Skills, Functional Communication, Hyperactivity, Aggression, Atypicality, and Attention Problems) and on the SRP (Anxiety, Somatization, Attitude to School, Attitude to Teachers, and Sensation Seeking;Reynold & Kamphaus, 2004). The use of gender specific norm groups has been discussed for other rating scales. Waschbusch and King (2006) examined mother and teacher behavioral ratings of 1,491 elementary school-aged students and found when using gender specific norms, a small group of female students were rated with higher than average scores in ADHD and ODD behaviors yet did not meet *DSM-IV* criteria and this was not found for male students. Based on the warning of Reynold and Kamphaus (2004) and findings of Waschbusch and King the use of gender specific norm groups for interpreting rating scale responses requires careful consideration.

Approximately half of the gender differences found in the current study were also found by the creators of the BASC-2 during the development and standardization of the rating scales. This finding indicates that the use of the separate gender group norms is not necessary and a combined gender norm group is the best option for the interpretation of scores on the BASC and other behavior rating scales. If gender is unnecessary for scoring and interpretation of the data, this would eliminate the need to provide gender when completing the form. As a result, this would increase appropriateness of using this measure with students who are gender nonconforming. This might include students who may be transitioning from one gender to another or who do not define their identity by

female or male. Ultimately this leads us away from depending on a gender binary in the fields of school psychology and education and moving towards the discussion of gender as a concept rather than category. This transition has been long underway. For example, in the past intelligence tests have been used to compare males' and females' performance but now only include genders to match the population of use in the development of the test (Drozdick, et al., 2012; Roid & Pomplun, 2012; Schrank & Wendling, 2012; Singer, et al., 2012; Wahlstrom, et al., 2012). The results of this study provide a strong argument that social, emotional, and behavioral instruments should follow intelligence tests toward an approach that recognizes that gender is not a significant variable in behavior.

An appeal was made in the field of school psychology to move beyond stereotyping individuals and groups through research of group differences (Merrell, Ervin, & Gimpel, 2006). It was emphasized that significant variability existed with groups and significant overlap existed between groups (Merrell, Ervin, & Gimpel). The current findings support these considerations. Gender stereotypes for behavioral, emotional, and social functioning were not endorsed by student, parent, and teacher on all subscale ratings.

Implications for school psychologists. School psychologists must remain aware of the expectation of genders, including gender stereotypes and bias. These impact consultation, referrals, and the assessment process as a whole. Because school psychologists work on teams, understanding team members' expectations for students based on gender is important. This will allow for insight when addressing concerns of students' functioning and potential gender bias or stereotyped expectations that may be

placed on students. Knowing others' perspectives on gender overall will increase the ability of the team to address and support the students' needs.

These types of implications were reported in the field of school psychology over 40 years ago. Engin, Leppaluoto, and Fodor (1973) posited:

> First and foremost, school psychologists should singly and collectively examine their operative value systems and determine in what respects the non-conscious sexist ideology of American society makes them vulnerable to perpetuating the sex role status quo. Reading widely in recent literature and actual participation in consciousness-raising groups should serve to facilitate further awareness of the issues and one's own attitudes in respect to the issues (pp. 6-7).

Similar approaches to addressing prejudice based on ethnicity or racism can be used to address gender based expectations and stereotypes. At that time, Engin et al. provided a list of action steps to take to address gender stereotyping and improve equality in the school system and the field of school psychology. As school psychologists it is not only important to address these in ourselves but in those with whom we work and collaborate.

Data-based problem solving and gender stereotypes. The best practice of data-based decision making endorsed by NASP creates opportunities to recognize and address gender stereotypes or bias. Steps of the data-based problem solving process include identifying the problem, analyzing the problem, developing and implementing an intervention, and evaluating the intervention (Merrell et al., 2006). As found in the BASC-2 student, parent, and teacher responses, the genders were perceived and rated similarly. However, when working with teachers in schools gender differences and

gendered expectations exist. Thus, incorporating the data-driven problem solving model allows for an individualized approach for addressing students', families', and school staff needs and concerns. The student can be understood as an individual student rather than in the context of a group.

Research implications. Hyde (2014) included two reasons for studying gender similarities and differences, including: (a) understanding how gender stereotypes of psychological differences potentially influence people's behavior and (b) the influence of such stereotypes in policy.

Behavioral, social, and emotional expectations based on gender exist across cultures and are embedded in social institutions, such as in schools (Kimmel, 2011). For example, in a preschool level early childhood college curriculum textbook a "nonsexist education" section cautioned:

> Still another aspect of helping children value their own sexuality has to do with recognizing the boys' needs for high-physical-energy activities and meeting their needs for role models in the children's center…However, much of the curriculum that might remediate deficiencies in little girls' education is left to "choice and chance." Participation in such activities as block play and large muscle activities that might also aid in developing spatial awareness are part of the self-select time (as is a selection of various science activities) (Hendrick & Weissman, 2006, pp. 241-244).

The authors of the textbook attempted to address issues of sexism and gender inequity while simultaneously promoting the reinforcement of gender differences in early education programs. This is an example of a textbook used in teacher training programs

for preschool students who are developing their gender identity. Future research might include investigating respondents' rating scale input followed by their perceptions of behavioral, social, and/or emotional functioning as related to gender.

The assumption of gender differences has policy and legal implications. For example, single sex classrooms have been thought to address gender equity and highlight gender differences in behavior, social, and learning needs (Arms, 2007; Hubbard & Datnow, 2005; Speilhagen, 2011; The State Education Resource Center, 2013). The reduction in Title IX limitations facilitated the creation and options for single-sex classrooms (Bailey, 2013; Hubbard & Datnow, 2005). Thus, a focus on gender differences permeated a ground breaking federal civil rights law that worked towards gender equity in schools. This demonstrates the power behind research in informing policy and highlights a reason to push for more research in gender similarities.

Examining disproportionality of female and male students in special education is another direction in which future studies might take. The number of female and male students receiving special education services, disability categories, and similar statistics can be researched. Additionally, future research might entail looking at the numbers of female and male students in relation to epidemiological studies and DSM-5 related categories.

Researchers might examine school psychologists' views and beliefs of gender in reference to social, emotional, and behavioral functioning of students. Gender-based stereotypes may be the foci of future studies that include the identification of internalizing and externalizing behaviors by school psychologists. Additionally, the referral process may be reviewed for how gender influences the referral of students for

additional support. Finally, the use of nondiscriminatory assessment and the problem-solving approach by school psychologists should be surveyed to examine practices in school settings.

In the current study the focus was on student, parent, and teacher ratings on the BASC-2. These included examining the full distribution which includes behavior levels that are subclinical or would not meet clinical criteria and are average and below. Whereas school psychologists are assessing for high levels or frequencies of behaviors and symptoms that would be outliers and considered extreme in relation to a full distribution. This is important to take into consideration as future research directions examine how students are identified as needing extra support.

Lastly, in the present study there was an assumption that gender bias was present in students', parents', and teachers' perspectives; yet, it is evident in many ratings that no to few differences were conveyed by students and parents and few were reported by teachers. Gender differences are highly endorsed in cultures and societies (Kimmel, 2011). Presently, it appears that when systematic observation data were gathered with rating scales appeared reduced some gender stereotypes and bias for adaptive skills and problem behaviors. Future research should include examining this reduction in gender bias on rating scales and ways to connect this to practice in schools for educators. Other types of psychoeducational assessment tools can be reviewed for similarities between the genders.

Strengths

The current study reviewed of the one of the most commonly used behavior rating scales (Merrill, 2009). Examining how genders are similarly rated on the BASC-2

creates a deeper understanding of the scale and how to interpret ratings in practice. Approaching research questions from a perspective of similarity provides another way to conceptualize research in school psychology. Tests of equivalence offer another perspective and way to view gender as a construct. This study brings voice to a discourse that has been quiet for 40 years in school psychology, gender issues. Lastly, investigating gender similarities draws focus to the inequity that exists between the sexes and the possible impacts this has on students' access to mental health supports in schools.

Limitations

Limitations with the current study exist in the statistical analysis and the behavioral assessment measure.

The BASC-2 rating scales. BASC-2 standardization and development did not include demographic information on the parent and teachers who completed the rating scales. It can be assumed more females completed the scale and this has impacts on the ratings and expectations of behavior. For example, in the 2011-2012 school year 76.3% of public school teachers were female (U.S. Department of Education Institute of Education Sciences National Center for Education Statistics, 2013). It is unclear how respondents' background and experiences influenced the ratings on the BASC-2. The BASC-2 rating scales are based on the perceptions of the raters and contexts which have may have impacted the ratings of behavior. Furthermore, in this study little gender bias was found in parent and teacher adaptive subscale ratings. Respondent information from the development of the BASC-2 could possibly inform the absence of gender bias on the majority of ratings and the gender differences found in various subscales for respondents.

Content validity and reliability are areas of limitation on the BASC-2 rating scales. Although the subscales were developed from factor analyses, not all subscales were aligned with diagnostic systems and correlated with narrow band measures. Temporal issues exist in the test-retest reliability estimates of the BASC-2 rating scales. The time frames at which test-retest administrations were conducted varied too much to interpret the strength of the reliabilities. Most of the responses to the rating scales are not observationally or objectively defined, leaving the responses open to the interpretation of the rater. The BASC-2 rating scales are perception based measures that are influenced by situations and should be interpreted in context. The context, situations, and differences between *never*, *sometimes*, *often*, and *almost always* responses impact the interpretation of student's positive skills and problem behaviors.

Tests of equivalence. Tests of equivalence are not commonly used, taught in most statistics courses, found in many statistics textbooks, nor readily available on the SPSS computer program (Weber & Popova, 2012). To complete tests of equivalence selecting intervals or effect sizes for interpretation is required. This has been argued both as a potentially arbitrary analysis (Ball et al., 2013) and as research based (Weber & Popova, 2012). In the current study no literature was found on which to base effect size interpretation and instead Cohen's d conservative guidelines were utilized. Thus, which measure is appropriate as a minimum substantial effect should be based on a meta-analysis of similar findings (Weber & Popova, 2012); however, this could not be found in the literature and is beyond the scope of this dissertation. An assumption of tests of equivalence is that the data examined be normally distributed. The data in the present study were not normally distributed for all subscales and this may have affected the

results. Thus, normality was violated, but there is no way to know how this might have impacted the current findings or if there is an impact on the results. Parametric statistics have been extensively used to develop and evaluate the BASC-2 (Reynolds & Kemphaus, 2004).

Conclusion

Until recently gender received little attention in the field of school psychology. Gender is viewed as a risk factor and issue in the diagnosis of many mental health disorders (American Psychiatric Association, 2013). The emphasis on gender differences has an impact on how students are referred and assessed for special education services. For example, male students have been found to be more likely to be identified with disabilities than female students (Coutinho & Oswald, 2005; Sullivan & Bal, 2013). It is important to consider how gender might impact students' access to supports in schools.

The purpose of this study was to examine the similarities between genders as perceived by students, parents, and teachers on a behavioral assessment measure. Strong similarity was found between the genders on all the subscales. The genders were not rated similarly on all the subscales at the moderate and small effect sizes. The relationships between students' gender and ratings were weak and little variation in the ratings were explained by gender. These findings indicate that few gender differences were rated by students, parents, and teachers on the BASC-2 moderate and strong effect sizes. Thus, despite the emphasis in the DSM-5 on gender differences for several mental health disorders, these differences were not endorsed in this study.

Best assessment practices promoted by the National Association of School Psychologists inherently avoid gender bias. Incorporating a problem solving model

approach forms an avenue by which to identify and address student concerns on an individual level without relying on what might be gender based stereotypes or gender based behavioral assumptions and standards. Non-discriminative assessment contributes another layer to the assessment process in which multiple data collecting methods and sources systematically acknowledge and recognize diversity in students. Employing these best assessment practices, school psychologists may take an active role in preventing the continuation of gender bias in special education referral and assessment processes. With this in mind the discourse on gender, which began over 40 years ago in school psychology, will continue and ideally progress towards action and equity in our schools.

References

Allen, J. L., Lavallee, K. L., Herren, C., Ruhe, K., & Schneider, S. (2010). DSM-IV criteria for childhood anxiety disorder: Informant, age, and sex differences. *Journal of Anxiety Disorders, 24*, 946-952. doi:10.1016/j.janxdis.2010.06.022

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Arlington, VA: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association. Retrieved from dsm.psychiatryonline.org

American Psychological Association (2006b). Think again: Women and men share cognitive skills. Retrieved from http://www.apa.org/research/action/share.aspx

American Psychological Association (2006b). Think again: Women and men share cognitive skills. Retrieved from http://www.apa.org/research/action/share.aspx

American Psychological Association, (2005a). Men: A different depression. Retrieved from http://www.apa.org/research/action/men.aspx

American Psychological Association, (2005a). Men: A different depression. Retrieved from http://www.apa.org/research/action/men.aspx

American Psychological Association, (2005b). Men and women: No big difference. Retrieved from http://www.apa.org/research/action/difference.aspx

American Psychological Association, (2005b). Men and women: No big difference. Retrieved from http://www.apa.org/research/action/difference.aspx

American Psychological Association, (2006a). Stereotype threat that widens achievement gap. Retrieved from http://www.apa.org/research/action/stereotype.aspx
American Psychological Association, (2006a). Stereotype threat that widens achievement gap. Retrieved from http://www.apa.org/research/action/stereotype.aspx
Arms, E. (2007). Gender equity in coeducational and single-sex environments. In Klein, S. K., Richardson, B., Grayson, D. A., Pollard, D. S., & Dwyer, C. A. (Eds.), *Handbook for Achieving Gender Equity through Education* (2nd ed., pp. 171-190). Mahwah, NJ: Lawrence Erlbaum Associates.

Bailey, S. M. (2013). Failing our kids: Despite pseudoscience to the contrary, sex segregation in public school creates problems – not solutions. *Ms. Magazine, Fall*. Retrieved from http://msmagazine.com/Fall2013/sexsegregation.asp

Ball, L. C., Cribbie, R. A., & Steele, J. R. (2013). Beyond gender differences: Using tests of equivalence to evaluate gender similarities. *Psychology of Women Quarterly, 37*(2), 147-154.

Bandura, A. (1971). *Social learning theory.* New York, NY: General Learning Press.

Bauermeister, J. J., Canino, G., Polanczyk, G., & Rohde, L. A. (2010). ADHD across cultures: Is there evidence for a bidimensional organization of symptoms? *Journal of Clinical Child & Adolescent Psychology, 39*(3), 362-372. doi:10.1080/15374411003691743

Bauermeister, J. J., Shrout, P. E., Chavez, L., Rubio-Stipec, M., Ramirez, R., Padilla, L., Anderson, A., Garcia, P., & Canino, G. (2007). ADHD and gender: Are risks and sequela of ADHD the same for boys and girls? *Journal of Child Psychology and Psychiatry, 48*(8), 831-839. doi:10.1111/j.1469-7610.2007.01750.x

Below, J. L., Skinner, C. H., Fearrington, J. Y., & Sorrell, C. A. (2010). Gender differences in early literacy: Analysis of kindergarten through fifth-grade Dynamic Indicators of Basic Early Literacy Skills probes. *School Psychology Review, 39*(2), 240-257.

Bem, S. L. (1978). Probing the promise of androgyny. In Sherman, J. A., & Denmark, F. (Eds.), *The psychology of women: New directions in research* (pp. 349-364). Madison, WI: Psychological Dimensions, Inc.

Berger, J. L., Addis, M. E., Reilly, E. D., Syzdek, M. R., & Green, J. D. (2012). Effects of gender, diagnostic labels, and causal theories on willingness to report symptoms of depression. *Journal of Social and Clinical Psychology, 31*(5), 439-457.

Boylan, K., Vaillancourt, T., Boyle, M., & Szatmari, P. (2007). Comorbidity of internalizing disorders in children with oppositional defiant disorder. *European Child and Adolescent Psychiatry, 16*, 484-494. doi:10.1007/s00787-007-0624-1

Bronfrenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32,* 513-530.

Brotman, M. A., Schmajuk, M., Rich, B. A., Dickstein, D. P., Guyer, A. E., Costello, E. J., Egger, H. L., Angold, A., Pine, D. S., & Leibenluft, E. (2006). Prevalence, clinical correlates, and longitudinal course of sever mood dysregulation in children. *Biological Psychiatry, 60*, 991-997. doi:10.1016/j.biopsych.2006.042

Bruchmüller, K., Margraf, J., & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis.

Journal of Counseling and Clinical Psychology, 80(1), 128-138. doi:10.1037/a0026582

Buss, D. M., & Schmitt, D. P. (2011). Evolutionary psychology and feminism. *Sex Roles, 64,* 768-787. doi: 10.1007/s11199-011-9987-3

Bussey, K., & Bandura, A. (1999). Social change theory of gender development and differentiation. *Psychological Review, 106*(4), 676-713.

Byrne Yates, C., Martinez, E., & Harrison, M. (2006). Gender-role nonconformity, suicide risk, and prevention. *National Association of School Psychologists Communiqúe Online, 35*(2).

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79*(5), 1185-1229.

Cattell, J. M. (1903). A statistical study of eminent men. *Popular Science Monthly, 62*, 359-377.

Chaplin, T. M., Gillham, J. E., & Seligman, M. E. P. (2009). Gender, anxiety, and depressive symptoms: A longitudinal study of early adolescents. *The Journal of Early Adolescence, 29*(2), 307-327. doi:10.1177/0272431608320125

Chrisler, J. C. (2013). Teaching about gender: Rewards and challenges. *Psychology of Men & Masculinity, 14*(3), 264-267. doi:10.1037/a0033260

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1). 155-159.

Coladarci, T., Cobb, C. D., Minium, E. W., & Clarke, R. B. (2008). *Fundamentals of Statistical Reasoning in Education* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Coles, E. K., Slavec, J., Bernstein, M., & Baroni, E. (2010). Exploring the gender gap in referrals for children with ADHD and other disruptive behavior disorders. *Journal of Attention Disorders, 16*(2), 101-108. doi:10.1177/1087054710381481

Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Craske, M. G., Kircanski, K., Phil, C., Epstein, A., Wittchen, H., Pine, D. S., Lewis-Fernandez, R., & Hinton, D. (2010). Panic disorder: A review of DSM-IV panic disorder and proposals for DSM-V. *Depression and Anxiety, 27*, 93-112. doi:10.1002/da.20654

Cribbie, R. A., Gruman, J. A., & Arpin-Cribbie, C. A. (2004). Recommendations for applying tests of equivalence. *Journal of Clinical Psychology, 60*(1), 1-10. doi:10.1002/jclp.10217

Cuffe, S. P., Moore, C. G., & McKeown, R. E. (2005). Prevalance and correlates of ADHD symptoms in the National Health Interview Survey. *Journal of Attention Disorders, 9*(2), 392-401. doi:10.1177/1087054705280413

D.M. (2009). What Larry Summers said- and didn't say. *Swarthmore College Bulletin: The Magazine of Swarthmore College*. Retrieved from http://bulletin.swarthmore.edu/bulletin-issue-archive/index.html%3Fp=145.html

Denmark, F. L., Klara, M., Baron, E. & Cambareri-Fernandez, L. (2008). Historical development of the psychology of women. In Denmark, F. L. & Paludi, M. A.

(Eds.), *Psychology of women: A handbook of issues and theories* (3-39). Westport, CT: Praeger.

Diamantopoulou, S., Henricsson, L., & Rydell, A., (2005). ADHD Symptoms and peer relations of children in a community sample: Examining associated problems, self-perceptions, and gender differences. *International Journal of Behavioral Development, 29*(5), 388-398. doi:10.1080/01650250500172756

Diamantopoulou, S., Verhulst, F. C., & van der Ende, J. (2011). The parallel development of ODD and CD symptoms from early childhood to adolescence. *European Child and Adolescent Psychiatry, 20,* 301-309. doi:10.1007/s00787-011-0175-3

Diseth, A., Meland, E., & Breidablik, H. J. (2014). Self-beliefs among students: Grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence. *Learning and Individual Differences, 35,* 1-8. doi:10.1016/j.lindif.2014.06.003

Dixon, A. L., Schneidegger, C., & McWhirter, J. J. (2009). The adolescent mattering experience: Gender variations in perceived mattering, anxiety, and depression. *Journal of Counseling & Development, 87,* 302-310.

Doll, B., & Cummings, J. A. (2008). *Transforming school mental health services: Population-based approaches to promoting the competency and wellness of children.* Thousand Oaks, CA: National Association of School Psychologists.

Domschke, K., Deckert, J., O'Donovan, M. C., & Glatt, S. J. (2007). Meta-analysis of COMT val158met in panic disorder: Ethnic heterogeneity and gender specificity.

American JOunral of Medical Genetics Part B (Neuropsychiatric Genetics), 144B, 667-673.

Driessen, G., & van Langen, A. (2013). Gender differences in primary and secondary educationL Are girls really out performing boys?. *International Review of Education, 59*,67-86. doi:10.1007/s11159-013-9352-6

Drozdick, L. W., Walhstrom, D., Zhu, J., & Weiss, L. G. (2012). The Wechsler Adult Intelligence Scale-Fourth Edition and the Wechsler Memory Scale-Fourth Edition. In D. P. Flanagan & P. L. Harrison (Eds.) *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 197-221). New York, NY: The Guilford Press.

DuPaul, G. J., Jitendra, A. K., Tresco, K. E., Vile Junod, R. E., Volpe, R. J., & Lutz, J. G. (2006). Children with Attention Deficit Hyperactivity Disorder: Are there gender differences in school functioning? *School Psychology Review, 35*(2), 292-308.

Eaton, A. A., & Rose, S. M. (2013). The application of biological, evolutionary, and sociocultural frameworks to issues of gender in introductory psychology textbooks. *Sex Roles, 69*, 536-542. doi:10.1007/s11199-013-0289-9

Elliott, S. N., & Busse, R. T. (1993). Behavior rating scales: Issues of use and development. *School Psychology Review, 22*(2).

Emery, E. J. (1973). The effects of sexism on the classification of children with learning disability. *The School Psychology Digest, 2*(3), 44-46.

Engin, A. W., Leppaluoto, J., & Fodor, I. (1973). Male and female- the mutually disadvantaged: The school psychologist's role in expanding options for both sexes. *The School Psychology Digest, 2*(3), 2-10.

Epstein, C. F. (1988). *Deceptive distinctions: Sex, gender, and the social order.* New Haven, CT: Yale University Press.

Fehm, L., Pelissolo, A., Furmack, T., & Wittchen, H. (2005). Size and burden of social phobia in Europe. *European Neuropsychopharmacology, 15*, 453-462. doi:10.1016/j.euroneuro.2005.04.002

Forestell, N. M., & Moynagh, M. (2014). General introduction. In N. M. Forestell & M. Moynagh, (In Eds.), *Documenting first wave feminisms* (pp. xvii-xxi). Toronto, Canada: Canadian Electronic Library.

Friedrich, A. A., Raffaele Mendez, L. M., & Mihalas, S. T. (2010). Gender as a factor in school-based mental health service delivery. *School Psychology Review, 39*(1), 122-136.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child Adolescent Psychiatry, 36*(8), 1036-1045.

Gerner, M. E., & Perry, F. (2000). Gender differences in cultural acceptance of career orientation among internationally mobile and non-internationally mobile adolescents. *School Psychology Review, 29*(2), 267-283.

Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders, 5,* 143-154. doi:10.1177/108705470200500302

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Graetz, B. W., Sawyer, M. G., & Baghurst, P. (2005). Gender differences among children with *DSM-IV* ADHD in Australia. *Journal of American Academy of Child & Adolescent Psychiatry, 44*(2), 159-168.

Haldeman, D. C. (2000). Gender atypical youth: Clinical and social issues. *School Psychology Review, 29*(2), 192-200.

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.

Hankin, B.L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology, 107*(1), 128-140.

Hanson, K., Smith, S. J. (2001). Gender equity in education: Change and challenge. In H. Rousso & M. L. Wehmeyer (Eds.), *Double jeopardy: Addressing gender equity in special education* (pp. 59-94). Albany, NY: State University of New York Press.

Hass, M. R., Brown, R. S., Brady, J., & Boehm Johnson, D. (2010). Validating the BASC-TRS for use with children and adolescents with an educational diagnosis of autism. *Remedial and Special Education, 20*(10), 1-12. doi:10.1177/0741932510383160

Hawkesworth, M. (1997). Confounding gender. *Signs: Journal of Women in Culture and Society, 22*(3), 649-685.

Hemel, D. J. (2009). Summers' comments on women and science draw ire: Remarks at a private conference stir criticism, media frenzy. *The Harvard Crimson.* Retrieved

from http://www.thecrimson.com/article/2005/1/14/summers-comments-on-women-and-science/

Hendrick, J. & Weissman, P. (2006). *The Whole Child* (8th ed.). Columbus, OH: Pearson Merrill Prentice Hall.

Hettema, J. M., Neale, M. C., & Kendler, K. S. (2001). A review and meta-analysis of genetic epidemiology of anxiety disorders. *American Journal of Psychiatry, 158*(10), 1568-1578.

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In Brooks-Gunn, J., & Petersen, A. (Eds.), *Girls at Puberty: Biological and Psychosocial Perspectives* (pp. 201-208). Hillsdale, NJ: :Lawrence Erlbaum Associates.

Hollingworth, L. S. (1914). Variability as related to sex differences in achievement: A critique. *American Journal of Sociology, 19*(4), 510-530.

Hubbard, L., & Datnow, A. (2005). Do single-sex schools improve the education of low-income minority students? An investigation of California′s public single-gender academies. *Anthropology and Education Quarterly, 36*(2), 115-131.

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65*, 373-398. doi:10.1146/annurev-psych-010213-115057

Hyde, J. S., & Plant, E. A. (1995). Magnitude of psychological gender differences: Another side to the story. *American Psychologist, 50*(3), 159-161.

Hyde, J. S., (1990). Meta-analysis and the psychology of gender differences. *Signs: Journal of Women in Culture and Society, 16*(1), 55-73.

Hyde, J. S., (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581-592.

Hyde, J. S., Mezulis, A. H., & Abramson, L. Y. (2008). The ABCs of depression: Integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. *Psychological Review, 115(2),* 291-313. doi:10.1037/0033-295X.115.2.291

Hyde, J. S., (2014). Gender similarities and differences. *Annual Review of Psychology, 65*, 373-398.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 20 U.S.C. § 618 (2004).

Kessler, et al., (2006). The prevalence and correlates of adult ADHD in the United States: Results from the National Comorbidity Survey Replication. *American Journal of Psychiatry, 163*(4), 716-723.

Kessler, R. C., Berglund, P., Demler, O., Robert, J., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E., & Wang, P. S. (2003). The epidemiology of major depressive disorder. *Journal of American Medical Association, 289*(23), 3095-3105.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of the 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 617-709.

Kessler, R. C., McGonagle, K. A., Swartz, M., Blazer, D. G., & Nelson, C. B. (1993). Sex and depression in the National Comorbidity Survey I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders, 29*, 85-96.

Kim, H. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52-54. doi:10.5395/rde.2013.38.1.52h

Kimmel, M. (2008). *The gendered society* (2nd ed.). New York, NY: Oxford University Press, Inc.

Kimmel, M. (2011). *The gendered society* (4th ed.). New York, NY: Oxford University Press, Inc.

Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In Maccoby, E. E. (Ed.), *The Development of Sex Differences* (pp. 82-173). Stanford, CA: Stanford University Press.

Kramer, M. D., Kruger, R. F., & Hicks, B. M. (2008). The role of internalizing and externalizing liability factors in according for gender differences in the prevalence of common psychopathological syndromes. *Psychology Medicine, 38*, 51-61. doi:10.1017-S0033291707001572

Krolokke, C., & Sorenson, A. S. (2006). *Gender communication theories and analyses.* Thousand Oaks, CA: Sage Publications, Inc.

Lahelma, E. (2014). Troubling discourse in gender and education. *Educational Research, 56*(2), 171-183. doi:10.1080/00131881.2014.898913

Lamiell, J. T. (1998). 'Nomothetic' and 'Idiographic': Contrasting Windelband's understanding with contemporary usage. *Theory & Psychology, 8*(1), 23-38. doi:10.1177/0959354398081002

Lavigne, J. V., LeBailly, S. A., Hopkins, J., Gouze, K. R., & Binns, H. J. (2009). The prevalence of ADHD, ODD, depression, anxiety in a community sample of 4-year-olds. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 315-328.

LeBleau, R. T., Glenn, D., Liao, B., Wittchen, H., Beesdo-Baum, K., Ollendick, T., & Craske, M. G. (2010). Specific phobia: A review of DSM-IV specific phobia and preliminary recommendations for DSM-V. *Depression and Anxiety, 27*, 148-167. doi:10.1002/da.20655

Lehto-Salo, P., Närhi, V., Ahonen, T., & Marttunen, M. (2009). Psychiatric comorbidity more common among adolescent females with CD/ODD than among males. *Nordic Journal of Psychiatry, 63*(4), 308-315. doi:10.1080/08039480902730615

Leibenluft, E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of the bipolar disorder in youths. *American Journal of Psychiatry, 168* (2), 129-142.

Levy, F, Hay, D. A., Bennett, K. S., &McStephen, M. (2004). Gender differences in ADHD subtypes comorbidity. *Journal of American Academy Child Adolescent Psychiatry, 44*(4), 368-376. doi:10.1097/01.chi.0000153232.64968

Liben, L. S., & Bigler, R. S., (2002). *The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways.* Boston, Massachusetts: Blackwell Publishing.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Marcotte, D., Alain, M., & Gosselin, M. (1999). Gender differences in adolescent depression: Gender-typed characteristics or problem-solving skills deficits? *Sex Roles, 41,* 31-48.

Maughan, B., Rowe, R., Messer, J., Goodman, R., & Meltzer, H. (2004). Conduct disorder and oppositional defiant disorder in a national sample: Developmental epidemiology. *Journal of Child Psychology and Psychiatry, 45*(3), 609-621.

Merrell, K. W. (2003). *Behavioral, Social, and Emotional Assessment of Children and Adolescents* (2nd ed.). Mahwah, NJ: Erlbaum.

Merrell, K. W. (2009). *Behavioral, Social, and Emotional Assessment of Children and Adolescents* (3rd ed.). New York, NY: Routledge.

Merrell, K. W. (2009). *Behavioral, social, emotional assessment of children and adolescents* (6th ed.). New York, NY: Routledge.

Merrell, K. W., & Walker, H. M. (2004). Deconstructing a definition: Social maladjustment versus emotional disturbance and moving the EBD field forward. *Psychology in the Schools, 41*(8), 899-910. doi:10.1002/pits.20046

Merrell, K. W., Ervin, R. A., & Gimpel, G. A. (2006). *School Psychology for the 21st Century: Foundations and Practices.* New York, NY: The Guilford Press.

Mischel, W. (1966). A social-learning view of sex differences in behavior. In Maccoby, E. E. (Ed.), *The Development of Sex Differences* (pp. 56-81). Stanford, CA: Stanford University Press.

Moffitt, T. E., Caspi, A., Rutter, M. & Silva, P. A., (2001). *Sex differences in antisocial behavior: Conduct disorder, delinquency, and violence in the Dunedin longitudinal study.* Cambridge, UK: Cambridge University Press.

Munkvold, L., Lundervold, A., Lie, S. A., & Manger, T. (2009). Should there be separate parent and teacher-based categories of ODD? Evidence from a general population. *The Journal of Child Psychology and Psychiatry, 50*(10), 1264-1272. doi:10.1111/j.14969-7610.2009.02091.x

Munkvold, L., Lundervold, A., Lie, S. A., & Manger, T. (2011). Oppositional defiant disorder: Gender differences in co-occurring symptoms of mental health problems in a general population of children. *Journal of Child Psychology, 39,* 577-587. doi:10.1007/s10802-011-9486-6

Murray-Close, D., & Crick, N. R. (2006). Mutual antipathy involvement: Gender and associations with aggression and victimization. *School Psychology Review, 35*(3), 472-492.

National Association of School Psychologists, (2004). NASP position statements: Racism, prejudice, and discrimination. In Thomas, A., & Grimes, J. (Eds.), *Best Practices in School Psychology V* (cxxiii-cxxiv). Bethesda, MD: National Association of School Psychologists.

National Association of School Psychologists, (2006). *School psychology: A blueprint for training and practice III.* Bethesda, MD: National Association of School Psychologists.

National Association of School Psychologists, (2008). *Degree type of practicing school psychologists by gender and ethnicity* (Unpublished data from the 2004-05 NASP Membership Survey). Retrieved from http://www.nasponline.org/standards/apamla_genderethnicity.aspx

National Association of School Psychologists, (2009). *Appropriate behavioral, social, and emotional supports to meet the needs of all students* (Position Statement). Bethesda, MD: National Association of School Psychologists.

Newman, J. L., Fuqua, D. R., Gray, E. A., & Simpson, D. B. (2006). Gender differences in the relationship of anger and depression in a clinical sample. *Journal of Counseling & Development, 84,* 157-162.

Nock, M. K., Kazdin, A., Hiripi, E., & Kessler, R. C. (2006). Prevalence, subtypes, and correlates of DSM-IV conduct disorder in the National Comorbidity Survey Replication. *Psychology Medicine, 36*, 699-710. doi:10.1017/S0033291706007082

Nock, M. K., Kazdin, A., Hiripi, E., & Kessler, R. C. (2007). Lifetime prevalence, correlates, and persistence of oppositional defiant disorder: Results from the National Comorbidity Survey Replication. *Journal of Child Psychology and Psychiatry, 48*(7), 703-713. doi:10.1111/j.1469-7610.2007.01733.x

Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., & Lipman, E. L. (1996). Integrating assessment data from multiple informants. *Journal of the Academy of Child and Adolescent Psychiatry, 35*(8), 1078-1085.

Ohan, J. L., & Visser, T. A. W. (2009). Why is there a gender gap in children presenting for attention deficit/hyperactivity disorder services? *Journal of Clinical Child & Adolescent Psychology, 3*(5), 650-660. doi:10.1080/15374410903103627

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25*, 241-286. doi:10.1006/ceps.20000.1040

Ortiz, S. O. (2008). Best practices in nondiscriminatory assessment. In Thomas, A., & Grimes, J. (Eds.), *Best Practices in School Psychology V* (pp. 661-678). Bethesda, MD: The National Association of School Psychologists.

Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The effects of single-sex compared with coeducational schooling on students′ performance and attitudes: A meta-analysis. *Psychological Bulletin, 3*, 1-31. doi:10.1037/a0035740

Polanczyk, G., Silva de Lima, M., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The Worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry, 164*, 942-948.

Priess, H. A., Lingberg, S. M., & Hyde, J. S. (2009). Adolescent gender-role identity and mental health: Gender intensification revisited. *Child Development, 80*(5), 1531-1544.

Reid, R., Riccio, C. A., Kessler, R. H., DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Rogers-Adkinson, D., & Noll, M. (2000). Gender and ethnic differences in ADHD as assessed by behavior ratings. *Journal of Emotional and Behavioral Disorders, 8*(1), 38-48.

Reilly, D., Neumann, D. L., & Andrews, G. (2014). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Education Progress assessments. *Journal of Educational Psychology*. Retrieved from http://dx.doi.org/10.1037/edu0000012

Reschly, D. J. (2008). School psychology paradigm shift and beyond. In Thomas, A., & Grimes, J. (Eds.), *Best Practices in School Psychology V* (pp. 3-13). Bethesda, MD: The National Association of School Psychologists.

Reynolds, C.R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children: Manual.* Circle Pines, MN: AGS Publishing.

Riglin, L., Petrides, K. V., Frederickson, N., & Rice, F. (2014). The relationship between emotional problems and subsequent school attainment: A meta-analysis. *Journal of Adolescence, 37*, 335-346. doi:10.1016/j.adolescence.2014.02.010

Roid, G. H., & Pomplum, M. (2012) The Stanford-Binet Intelligence Scales, Fifth Edition. In D. P. Flanagan & P. L. Harrison (Eds.) *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 249-268). New York, NY: The Guilford Press.

Rousso, H., & Wehmeyer, M. L. (2001). Introduction. In H. Rousso & M. L. Wehmeyer (Eds.), *Double jeopardy: Addressing gender equity in special education* (pp. 1-9). Albany, NY: State University of New York Press.

Rucklidge, J. J. (2006). Gender differences in neuropsychological functioning of New Zealand adolescents with and without attention deficit hyperactivity disorder. *International Journal of Disability, Development and Education, 53*(1), 47-66. doi:10.1080/10349120600577402

Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the USA: Results from the National Comorbidity Survey Replication. *Psychological Medicine, 38*, 15-28. doi:10.1017/S0033291707001699

Rushton, J. L., Forcier, M., & Schectman, R. M. (2002). Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(2), 199-205.

Rydell, A. M. (2010). Family factors and children's disruptive behavior: An investigation of links between demographic characteristics, negative life events and symptoms of ODD and ADHD. *Social Psychiatric Epidemiology, 45,* 233-244. doi:10.1007/s00127-009-0060-2

Salkind, N. J. (2008). *Statistics for people who think they hate statistics* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Satler, J. A. (2002). *Assessment of children: Behavioral and clinical applications* (4th ed.). La Mesa, CA: Jerome M. Satler, Publisher, Inc.

Shrank, F. A., & Wendling, B. J. (2012). The Woodcock-Johnson III Normative Update: Tests of Cognitive Abilities and Tests of Achievement. In D. P. Flanagan & P. L. Harrison (Eds.) *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 297-335). New York, NY: The Guilford Press.

Schuirmann, D. J. (1987). Pharmacometrics: A comparison of the two-one sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics, 15*(6).

Sciutto, M. J., Nolfi, C. J., & Bluhm, C. (2004). Effects of child gender on symptom type on referrals for ADHD by elementary school teachers. *Journal of Emotional and Behavioral Disorders, 12(4),* 247-253.

Seedat, S., et al. (2009). Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry, 66*(7), 785-795.

Sherif, C. W. (1979). Bias in psychology. In Sherman, J. A., & Beck, E. T. (Eds.), *The Prism of Sex: Essays in the Sociology of Knowledge* (pp. 93-133). Madison, WI: The University of Wisconsin Press.

Shields, S. A. (1975). Functionalism, Darwinism, and the psychology of women: A study in social myth. *American Psychologist,* 739-754.

Singer, J. K., Lichtenberger, E. O., Kaufman, J. C., Kaufman, A. S., & Kaufman, N. L. (2012). The Kaufman Assessment Battery for Children-Second Edition and the Kaufman Test of Educational Achievement-Second Edition. In D. P. Flanagan & P. L. Harrison (Eds.) *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 269-296). New York, NY: The Guilford Press.

Spielhagen, F. R. (2011). "It all depends…": Middle school teachers evaluate single-sex classes. *Research in middle level education, 34*(7), 1-12.

State Education Resource Center (2013). *Single-sex education: The Connecticut context technical report*. Middletown, CT: State Education Resource Center.

Steffens, M. C., Jelenec, P., & Noack, P., (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology, 102*(4), 947-963. doi:10.1037/a0019920

Thomas, A., & Grimes, J. (Eds.) (2008). *Best Practices in School Psychology V*. Bethesda, MD: The National Association of School Psychologists.

Thompson Woolley, H. (1910). A review of the recent literature on the psychology of sex. *Psychological Bulletin, 7*(18), 335-342.

Thorndike, E. L. (1908). Sex in education. *The Bookman, 23*, 211-214.

Thorndike, E. L. (1910). *Educational psychology* (2nd ed.). New York, NY: Teachers College, Columbia University.

Trew, K., & Kremer, J. (Eds.) . (1998). *Gender and psychology.* London, New York: Arnold.

Turk, C. L., & Heimburg, R. G. (1998). An investigation of gender differneces in social phobia. *Journal of Anxiety Disorders, 12*(3), 209-223.

U.S. Department of Education. (n.d.). Building the legacy: IDEA 2004 [Search results: Query gender.] Retrieved from http://idea.ed.gov/explore/search/GO_x/0/GO_y/0/query/gender/search_option/all

U.S. Department of Education Institute of Education Sciences National Center for Educational Statistics (2013). *Table 209.10 Number and percentage of distribution of teachers in public and private elementary and secondary schools, by selected teacher characteristics: Selected years, 1987-1988 through 2011-2012.* Retrieved from htt://nces.ed.gov/programs/digest/d13/tables/dt13_209.10.asp

Uddin, M., Koenen, K. C., de los Santos, R., Bakshis, E., Aiello, A. E., & Galea, S. (2010). Gender differences in the genetic and environmental determinants of adolescent depression. *Depression and Anxiety, 27,* 658-666. doi:10.1002/da.20692

Vecchione, M., Alessandri, G., & Marsicano, G. (2014). Academic motivation predicts educational attainment: Does gender make a difference?. *Learning and Individual Differences, 32*, 124-131. doi:10.1016/j/lindif.2014.01.003

Vesga-Lopez, O., Schneier, F. R., Wang, S., Heimburg, G., Liu, S., Hasin, D. S., & Blanco, C. (2008). Gender differences in generalized anxiety disorder: Results from the National Epidiologic Survey on Alcohol and Related Conditions (NESARC). *Journal of Clinical Psychiatry, 69*, 1606-1616.

Wade, T. J., Cairney, J., & Pevalin, D. J. (2002). Emergence of gender differences in depression during adolescence: National panel results from three countries. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(2), 190-198.

Wahlstrom, D., Breaux, K. C., Zhu, J., & Weiss, L. G. (2012). The Wechsler Preschool and Primary Scale of Intelligence-Third Edition, the Wechsler Intelligence Scale for Children-Fourth Edition, and the Wechsler Individual Achievement Test-Third Edition. In D. P. Flanagan & P. L. Harrison (Eds.) *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 224-248). New York, NY: The Guilford Press.

Waschbusch, D. A., & King, S. (2006). Should sex-specific norms be used to assess attention-deficit/hyperactivity disorder or oppositional defiant disorder? *Journal of Consulting and Clinical Psychology, 74*(1), 179-185. doi:10.1037/0022-006X.74.1.179

Weber, R., & Popova, L. (2012). Testing Equivalence in communication research: Theory and application. *Communication Methods and Measures, 6*, 190-213. doi:10.1080/19312458.2012.703834

Westen, D. (2002). *Psychology: Brain, Behavior, and Culture* (3rd ed.). New York, NY: John Wiley & Sons, Inc.

Wilson, M. S., & Reschly, D. J., (1995). Gender and school psychology: Issues, questions, and answers. *School Psychology Review, 24*(1), 45-61.

Witchen, H., Gloster, A. T., Beesdo-Baum, K. Fava, G. A., & Craske, M. G. (2010). Agoraphobia: A review of the diagnostic classificatory position and criteria. *Depression and Anxiety, 27*, 113-133. doi:10.1002/da.20646

Wittchen, H., Stein, M. B., & Kessler, R. C. (1999). Social fears and social phobia in a community sample of adolescents and young adults: Prevalence, risk factors and co-morbidity. *Psychological Medicine, 29*, 309-323.

Young, E.L., Sabbah, H.Y., Young, B. J., Reiser, M. L., & Richardson, M. J. (2010). Gender differences and similarities in a screening process for emotional and behavioral risks in secondary schools. *Journal of Emotional and Behavioral Disorders, 18*(4), 225-235. doi:10.1177/1063426609338858

Appendices

Appendix A

Internalizing Disorders: Mood Disorders and Anxiety Disorders

Study	Study Design	Participants	Disorder(s) Cited for in the DSM-5	Measure or Inclusion Criteria for Meta-Analyses	DSM-5 Citation	Gender Reference(s)
Wittchen et al. (2010)	Literature Review	469 Articles	Agoraphobia	469 relevant citations	Females are twice as likely as males to experience agoraphobia	A stronger female preponderance was found for agoraphobia without panic than for panic disorder (p. 119), as based on two studies. To our knowledge, there are no studies that have reported significant age x gender interaction differences between panic disorder and agoraphobia (p. 119)
Brotman et al. (2006)	Longitudinal Epidemiological Study	N=1,420 children, ages 9-19-years-old Western North Carolina	Disruptive Mood Dysregulation Disorder	Items from the Child and Adolescent Psychiatric Assessment (CAPA) parent and child-based interviews	Among community samples, a male preponderance appears to be supported	Severe mood dysregulation (SMD) a clinical syndrome characterized by extreme, impairing, and chronic irritability with hyperarousal symptoms. 77.6% of participants meeting SMD criteria were male and 66.3% of those meeting severe impairment were male (p. 994)
Leibenluft (2011)	Longitudinal Study; Cross-Sectional Studies of Family History and	N=84 youths with SMD; 93 youths with DSM-IV Bipolar Disorder	Disruptive Mood Dysregulation Disorder	Rates of mood episode Parent diagnoses;	Children presenting to clinics with features of disruptiv	96 male of 111 participants studied at the National Institute of Mental Health

	Pathophysiology	N=39 youths with SMD or Bipolar Disorder		functional MRI; the Child Behavior Checklist	e mood dysregulation disorder are predominantly male	from 2002 through 2011 approximately
Seedat et al. (2009)	Face-to-Face Household Surveys	N=72,933; GAD N=15 Colombia, Lebanon, Mexico, South Africa, Ukraine, Belgium, France, Germany, Israel, Italy, Japan, the Netherlands, New Zealand, Spain, the United States	Generalized Anxiety Disorder (GAD)	2-part interviews; the World Health Organization Composite International Diagnostic Interview (CIDI)	Females are twice as likely as males to experience generalized anxiety disorder	Results are highly consistent across countries in showing that women have a significantly higher lifetime risk of most mood disorders (MDD and dysthymic disorder) and all anxiety disorders than men (p. 789). Ratio of female to male is 1.7 with 95% confidence interval 1.5 to 1.9 (p. 789)
Vesga-Lopez et al. (2008)	Cross-sectional survey of the U.S. population	N=43,093	Generalized Anxiety Disorder	The National Institute on Alcohol Abuse and Alcoholism's (NIAAA's) 2001-2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC); The Alcohol Use Disorder and Associated Disability Interview Schedule-DSM-IV	Females are twice as likely as males to experience generalized anxiety disorder In females, comorbidity is largely confined to the anxiety disorders and unipolar depression, whereas in males, comorbidity is more likely to extend to the substance use disorders as well	The lifetime prevalence rates of DSM-IV GAD were 2.8% for men and 5.3% for women, whereas the 12-month rates of GAD were 1.2% for men and 2.7% for women (p.1609) Men with 12-month lifetime GAD were significantly more likely than women to have any substance use disorder, any alcohol use disorder, drug dependence, and antisocial personality disorder (p. 1609)
Kramer, Krueger, and	Survey data via mail	N=2,992 middle aged twins	Generalized Anxiety Disorder	The Psychiatric Diagnostic Screening	In females, comorbidity is largely	The best-fitting two-factor strong

Hicks (2007)		Minnesota Twin Registry		Questionnaire (PDSQ); an adult antisocial behavior scale (the Bayesian information criterion (BIC))	confined to the anxiety disorders and unipolar depression, whereas in males, comorbidity is more likely to extend to the substance use disorders as well	invariance model differentiated internalizing (i.e. generalized anxiety disorder, major depressive disorder, panic disorder, social phobia, hypochondriasis, OCD, bulimia/binge-eating disorder, and agoraphobia) from externalizing (i.e. alcohol abuse/dependence, adult antisocial behavior, and drug abuse/dependence) syndromes in men and women indicated that the magnitude of the relationships between the latent factors and the observed syndromes (i.e. loadings) were equivalent across gender, and revealed that the differences in means of the latent factors of internalizing and externalizing. Compared to men, women exhibited greater mean levels of variability on the externalizing factor (p. 55)

Kessler et al. (2003)	Face-to face household survey	N=9,090; 18 years or older United States	Major Depressive Disorder	The World Health Organization's (WHO) Composite International Diagnostic Interview (CIDI); the Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR); the Sheehan Disability Scale (SDS), and the WHO Disability Assessment Scale (WHO-DAS), the Structured Clinical Interview for DSM-IV	Females experience 1.5-3-fold higher rates than males beginning in early adolescence	Major depressive disorder odds ratio, 95% confidence interval; females lifetime 1.7 (1.5-2.0), 12-month among lifetime 1.4 (1.1-1.8), 12-month severe among 12-month 1.3 (1.0-1.9) Males lifetime 1.0, 12-month among lifetime: 1.0, 12-month severe among 12-month: 1.0 (p. 3100)
Kessler, Chiu, Demler, and Walters (2005)	National representative face-to-face survey (epidemiological survey); structured interview	Part 1: N=9,282 Part 2: N=5,692 18 or older; English speaking United States	Panic Attack Specifier	World Mental Health (WMH) Composite International Diagnostic Interview (CIDI); the US National Comorbidity Survey Replication (NCS-R)	Females are more frequently affected than males, although this gender difference is more pronounced for panic disorder	Correlates of pure internalizing disorders (class 2) include being female and married, having a college education, and residing in the suburbs of small metropolitan areas (p. 622). The results regarding sociodemographic correlates are broadly consistent with precious surveys in finding that mental disorders (ie, low probability of membership in latent class 1) are associated with a general pattern of disadvantaged social status, including being female, unmarried, and having low socioeconomic

						status (p. 625). Sociodemographic correlates (odds ratios) of internalizing disorders, 95% confidence interval; females 1.6 (1.4-1.8); males 1.0; significant at the $P<0.5$ level, 2 sided test (p. 623)
Kessler, Chiu, Demler, and Walters (2005)	National representative face-to-face survey (epidemiological survey); structured interview	Part 1: N=9,282 Part 2: N=5,692 18 or older; English speaking United States	Panic Disorder (PD)	World Mental Health (WMH) Composite International Diagnostic Interview (CIDI); the US National Comorbidity Survey Replication (NCS-R)	Females are more frequently affected than males, at a rate of approximately 2:1 (p. 210)	Correlates of pure internalizing disorders (class 2) include being female and married, having a college education, and residing in the suburbs of small metropolitan areas (p. 622). The results regarding sociodemographic correlates are broadly consistent with precious surveys in finding that mental disorders (ie, low probability of membership in latent class 1) are associated with a general pattern of disadvantaged social status, including being female, unmarried, and having low socioeconomic status (p. 625). Sociodemographic correlates (odds ratios) of internalizing disorders, 95% confidence

						interval; females 1.6 (1.4-1.8); males 1.0; significant at the $P<0.5$ level, 2 sided test (p. 623)
Craske et al. (2010)	Literature review	DSM-IV review	Panic Disorder (PD)	Data published on or after 1994	The rates of panic disorder show a gradual increase during adolescence, particularly in females, and possibly following the onset of puberty, and peak during adulthood (p. 210)	The rate of PD shows a gradual increase during adolescence, particularly in girls, and possibly following the onset of puberty. Although the main gender differentiation occurs in adolescence, the gender difference is already observable before the age of 14 (p.106).
Domschke et al. (2007)	Meta-analysis	6 controlled case studies N=557 patients with Panic Disorder; Males=209, Females=319 N=763 patients in control group	Panic Disorder (PD)	Case-control studies of the functional val158metpolymorphism of the COMT gene in samples of clinically diagnosed participants with Panic Disorder	There is some evidence for sexual dimorphism, with an association between panic disorder and the catechol-O-methyltransferase (COMT) gene in females only (p. 212)	Five of the studies were stratified by gender…Analyses of the studies for males and females separately revealed no significant effect. However, in the female subgroup there was evidence for significant heterogeneity. Thus, the analysis was re-run for the female subgroup stratified for ancestry. This revealed a significant association between the COMT 158val allele with panic disorder in the

						Caucasian samples, but an opposite effect in Asian samples (pp. 668-669)
Hettema et al. (2001)	Meta-analysis	Family and twin studies of panic disorder, generalized anxiety disorder, phobias, and obsessive compulsive disorder (OCD) Panic Disorder: N=8 studies	Panic Disorder (PD)	Studies of adult participants that estimated the risk in relatives for the same anxiety disorder as diagnosed in the proband	There is some evidence for sexual dimorphism, with an association between panic disorder and the catechol-O-methyltransferase (COMT) gene in females only (p. 212)	Panic Disorder: The two larger twin studies, performed on different samples and opposite genders, are consistent with each other in attributing 30%-40% of the variance in liability comes from individual-specific environments (p. 1570) Generalized Anxiety Disorder: The best-fitting model predicted that 31.6% (95% CI=24%-39%) of variance for liability to generalized anxiety disorder was attributable to additive genetics in both genders and that the same genes predispose men and women to generalized anxiety disorder (p. 1571)
Allen et al. (2010)	Structured diagnostic interviews	N=106 total; 43 with primary separation anxiety disorder diagnosis 4-15 years-old	Separation Anxiety Disorder	The Diagnostic Interview for Children and Youth for DSM-IV-TR: Child and Parent Versions (Kinder-DIPS; DSM-IV-TR Version)	Girls manifest greater reluctance to attend school than boys. Indirect expression of fear of separation may be more common in males than in females, for example, by	Main effects of sex and age, and sex by age and informant by age interactions were all non-significant. A trend-level interaction between informant and sex emerged, F(1,

					limited independent activity, reluctance to be away from home alone, or distress when spouse do things independently or when contact with spouse or offspring is not possible (p. 193)	134.10)=3.01, p=.085, *d*=.30. A follow-up *t*-test examining number of symptoms by sex separately for parent and child reports indicated that parents tended to report more symptoms for girls (mean=4.93, SD=1.53) than for boys (mean=4.29, SD=1.59); *t*(104)=-2.11, *p*<.05, *d*=.41. While statistically significant, with a medium effect, this difference is not clinically significant, meaning there is no difference between 4.93 and 4.29 symptoms in meeting diagnostic criteria, as the threshold is set at three symptoms (p. 949) Girls and younger children were more likely to manifest reluctance or refusal to go to school or elsewhere because of fear of separation (950) We found only a trend-level difference of sex on the overall number of symptoms reported (p. 950)
Fehm,	Literature	Epidemiologi	Social	Studies published	In general,	Prevalence

Pelissolo, Furmark, and Wittchen (2005)	review	cal studies in European countries, Iceland, Norway, and Switzerland	Anxiety Disorder (Social Phobia)	after 1980 with DSM-III and later diagnostic criteria	higher rates of social anxiety disorder are found in females than males in the general population (with odds ratios ranging from 1.5 to 2.2) (p. 204) Gender rates are equivalent or slightly higher for males in clinical samples, and it is assumed that gender roles and social expectations play a significant role in explaining the heightened help-seeking behavior in male patients (p. 204)	estimates were generally higher in women than in men (p.454) Consistent with studies from other regions of the world, gender differences have been documented, with women being more frequently affected by social phobia than men. Odds ratios range between 1.5 to 2.2. There is an ongoing discussion whether gender difference in social phobia is less pronounced than in other anxiety disorders. In clinical samples, usually no gender differences or even slightly higher rates for men have been reported, and it is assumed that gender roles and social expectations play a significant role in explaining the heightened help-seeking behavior found in male patients. Other sociodemographic correlates are poor financial situation, low social class, single or unmarried status,

						unemployment and poor education, but it remains unclear whether those factors are antecedents or consequence of social phobia (p. 456)
Ruscio et al. (2008)	National representative face-to-face survey (epidemiological survey); structured interview	Part 1: N=9,282 Part 2: N=5,692	Social Anxiety Disorder (Social Phobia)	The World Health Organization (WHO) Composite International Diagnostic Interview Version 3.0 (CIDI 3.0); The National Comorbidity Survey Replication (NCS-R)	Females with social anxiety disorder report a greater number of social fears and comorbid depressive, bipolar, and anxiety disorders, whereas males are more likely to fear dating, have oppositional defiant disorder or conduct disorder, and use alcohol and illicit drugs to relieve symptoms of the disorder. Paruresis is more common in males (p. 206)	Social phobia involving 1-4 fears is more common among males and those of 'other' race-ethnicity (mostly American Indian or Asian). By contrast, social phobia involving a larger number of fears is significantly related to being younger, female, neither Hispanic nor non-Hispanic Black, never or previously married, neither a student nor retired, having less than a college education, an 'other' employment status, and low income (p. 20)
Turk et al. (1998)	Structured clinical interview	N=108 male, 104 female with principal diagnosis of social phobia	Social Anxiety Disorder (Social Phobia)	The Anxiety Disorders Interview Schedule-Revised (ADIS-R); The Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L); The Schedule of Affective Disorders and Schizophrenia-Lifetime for Anxiety Disorders (SADS-LA); The Structured Clinical Interview for DSM-III-R (SCID); The Social	Females with social anxiety disorder report a greater number of social fears and comorbid depressive, bipolar, and anxiety disorders, whereas males are more likely to fear dating, have oppositional defiant disorder or conduct	Men and women were equally likely to receive a comorbid diagnosis of avoidant personality disorder (APD) (44.4% vs. 44.2%). Furthermore, no gender differences were found when all three levels of APD (absent,

				Avoidance and Distress Scale (SADS); The Fear of Negative Evaluation Scale (FNE); The Social Phobia Subscale of Fear Questionnaire (FQ-Social); The Social Interaction Anxiety Scale (SIAS); The Social Phobia Scale (SPS); The Personality Disorder Examination (PDE); The Liebowitz Social Anxiety Scale (LSAS); 4 minute behavior test; Subjective Units of Discomfort Scale (SUDS)	disorder, and use alcohol and illicit drugs to relieve symptoms of the disorder. Paruresis is more common in males (p. 206)	probable, definite) were examined (p. 216) Men and women reported that suffering from social phobia for similar lengths of time (19.3 vs 20.3 years). No significant differences were observed in the proportion of men and women who reported previous psychotherapy (61.6% vs. 63.6%) or treatment with anxiolytic medication (38.4% vs. 31.8%). Men and women were equally likely to receive a comorbid diagnosis of a mood disorder (21.9% vs. 27.3%), an additional anxiety disorder (38.4% vs. 48.5%), or any mood or anxiety disorder (47.9% vs. 56.1%). Furthermore, no significant gender differences were observed in comorbidity rates for any specific mood or anxiety disorder (pp. 216-217) Upon further statistical analysis the only significant differences that

						emerged were that more women (96.2%) than men (86.1%) reported fear of going to a party and more men (38.0%) than women (21.2%) reported fear of urinating in a public restroom…Follow-up tests revealed…the only significant difference that emerged was that more men (30.6%) than women (14.4%) reported avoidance of urinating in a public restroom (p. 219)
Wittchen, Stein, and Kessler (1999)	Survey; follow-up surveys at 15 and 30-month intervals	N=3,021 14-24 year-old Munich, Germany	Social Anxiety Disorder (Social Phobia)	The Early Developmental Stages of Psychopathology (EDSP); The Computer-Assisted Personal Interview (CAPI) version of the Munich-Composite International Diagnostic Interview (M-CIDI); The Retrospective Self-Report of Inhibition Questionnaire (RSRI)	The gender difference is more pronounced in adolescents and young adults (p. 204)	Overall, 4.9% of males and 9.5% of females aged 14–24 meet lifetime DSM-IV criteria for social phobia. Overall, females have significantly higher lifetime and cross-sectional rates than males (P<0.001). The persistence of generalized social phobia (males, 76.8%; females, 90%) from lifetime (LT) into the past 12-months is slightly higher than in isolated social phobia (males 59.5%; females 69.2; P=0.03) (p. 313) A persistent strong fear of doing things in front of other

						people or being in the centre of attention at some time in their life, was acknowledged by 22.3% of male and 32.2% of female respondents. Women were significantly more likely to report these fears than men and were slightly more likely to report multiple fears than men (P<0.04) (p. 315) Most social phobias among respondents in the age range of the sample start between ages 10 and 21. But the generalized type has a significantly lower median age of onset (11.5 v. 14 for males, P<0.004 and 12.5 v. 15 for females, P<0.003) than nongeneralized social phobia (p. 318)
LeBleau et al. (2010)	Literature review	24 relevant papers	Academic published findings between 1994 and 2009	Specific Phobia (SP)	Females are more frequently affected than males, at a rate of approximately 2:1, although rates vary across different phobic stimuli. That is, animal, natural environment, and situational spec ific phobias are predominantly experienced by females,	Overall, women have higher prevalence rates of SP than do men. A study of a large Swedish sample found that 26.5% of all women and 12.4% of all men met criteria for a SP. In terms of types, animal phobia is more prevalent among women (4.3–12.1%) than men (2.7–3.3%). The

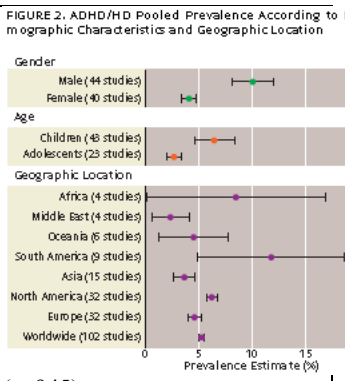
					whereas blood-injection-injury phobia is experienced nearly equally by both genders (p. 199)	same is true for natural environment phobia (men53.2%; women55.3%. Height phobia was the most common type of SP among men (3.3–6.3%), but the second or third most common type of SP among women (1.6–8.6%). Situational phobia is more prevalent among women (6.4–17.4%) than men (1.6–8.5%). Findings regarding the sex ratio of B-I-I phobia have been mixed, with two studies finding higher prevalence in females (female: 4.4–6.4%; male: 1.8–3.9%) and another finding no gender differences for prevalence rates. In sum, animal phobia, natural environment (height) phobia, and situational phobia all show higher prevalence among women than men whereas the findings regarding B-I-I phobia are mixed. Again, however, absolute differences in proportions across gender

						should be viewed with caution given the differences in ways of establishing impairment for phobias across studies (pp. 150-151)

Appendix B

Externalizing Disorders: Attention Deficit Hyperactivity Disorder, Conduct Disorder, and Oppositional Defiance Disorder

Study	Study Design	Participants	Disorder(s) Cited for in the DSM-5	Measure or Inclusion Criteria for Meta-Analyses	DSM-5 Citation	Gender Reference(s)
Kessler et al. (2006)	Retrospective assessment of childhood ADHD; semi-structured interview, clinical reappraisal interviews	N=3,199 18-44 year-old surveyed N=154 blinded clinical follow-up interviews	Attention Deficit Hyperactivity Disorder (ADHD)	The National Comorbidity Survey Replication; The Adult ADHD Clinical Diagnostic Scale Version 1.2; The World Health Organization (WHO) Composite International Diagnostic Interview (CIDI) Version 3.0 (23)	ADHD is more frequent in males than in females in the general population, with a ratio of approximately 2:1 in children) and 1.6:1 in adults. Females are more likely than males to present primarily with inattentive features (p. 63)	The multiple-imputation estimates of clinician-assessed adult ADHD were significantly higher among men, non-Hispanic whites (i.e., non-Hispanic blacks and Hispanics had significantly lower odds than non-Hispanic whites), the previously married, and people in the "other" employment category (mostly the unemployed and disabled). The odds ratios for these predictors were all modest in substantive terms (1.6–3.3)… A significantly higher proportion of women than men with adult ADHD had received treatment for mental or substance-related problems in the 12 months before the interview (53.1% versus 36.5%, z=2.6, p=0.02). However, only 25.2% of the treated respondents had received treatment for ADHD (22.8% of women and 27.7% of men, z= 0.5, p=0.60). Because of this low proportion, only 10.9% of the respondents with adult ADHD had received treatment for ADHD in the 12 months before interview (12.1% of women and 10.1% of men, z=0.4, p=0.66) (p. 718)
Polanczyk et al. (2007)	Literature review	9,105 records; 303 full-text articles N=171,756 participants worldwide	Attention Deficit Hyperactivity Disorder	Textbooks, articles, contact with authors and experts on ADHD/HD epidemiology Articles published between 1978-2005 in English, German,	ADHD is more frequent in males than in females in the general population, with a ratio of approximately 2:1 in children) and 1.6:1 in adults.	Age and gender were not included in the final multivariate metaregression model because less than 50% of the studies reported findings stratified by these variables, but individual estimates were computed according to these strata (age was stratified into the following ranges: 6–11 and 12–18) (p. 944)

				French, Spanish, and Portuguese	Females are more likely than males to present primarily with inattentive features (p. 63)	 (p. 945)
Card, Stucky, Sawalani, and Little (2008)	Meta-analysis	N=148 studies on child and adolescent direct and indirect aggression N=73,498 children	Conduct Disorder	Direct and indirect aggression; children younger than 18-years-old; normative samples	Whereas males tend to exhibit both physical aggression and relational aggression (behavior that harms social relationships of others), females tend to exhibit relatively more relational aggression (p. 474)	Overall, boys were more directly aggressive than girls, with an average effect that was medium in magnitude, .29. To facilitate comparison to other meta-analyses of gender differences, this effect size is equivalent to d=.61. First, method of assessing direct aggression is related to the magnitude of gender differences, with parent reports and self-reports yielding the smallest gender differences, whereas peer reports (nominations or ratings) and observations yield the largest differences. The percentage of ethnic minorities in samples also relates to the magnitude of gender differences in direct aggression, such that smaller differences are found with increasing percentages of ethnic minorities. In order to compare gender differences in physical versus verbal aspects of direct aggression, we conducted post hoc analyses of 27 studies that reported gender differences separately for these two aspects. Results indicate that gender differences are stronger for physical (effect size .34, equivalent d=.73) than verbal (effect size .19, equivalent d=.38) expressions of direct aggression…our (random effects) results indicate a negligible, but statistically different from zero, average gender difference, with girls exhibiting more indirect aggression than boys, effect size -.03 (equivalent d= -.06). As with direct aggression, this gender difference in indirect aggression varies by reporter.

						Specifically, parent and teacher reports yield gender differences of girls being higher than boys, whereas self-reports yield a slightly higher level for boys than girls. Other reports are not significantly different from zero. However, the magnitude of these gender differences in indirect aggression is trivial regardless of reporter (pp. 1193-1194)
Maughan, Rowe, Messer, Goodman, and Meltzer (2004)	Parent interview, child interview, teacher questionnaire	N=10,438; 5,212 boys, 5,226 girls United Kingdom	Conduct Disorder (CD)	The Development and Well-Being Assessment (DWBA)	Females with a diagnosis of conduct disorder are more likely to exhibit lying, truancy, running away, substance use, and prostitution (p. 474)	CD was significantly more common in boys than in girls (OR = 2.7, 95% CI=1.9–3.8, p < .001); across the full sample 42 girls (.8%) and 110 boys (2.1%) met DSM-IV criteria for CD. Among the youngest children in the sample, rates of CD were low in both genders. For boys risks then appeared to increase steadily with age, while in girls rates remained low until the early teens. To highlight possible developmental trends, figures for the boys are grouped into four age bands (ages 5–7 years, 8–10 years, 11–12 years and 13–15 years); because of the low prevalence of CD in girls before the teens, data for girls are only shown for the 13–15-year age-range. Among boys, total CD symptoms increased with age (POR=1.2, 95% CI=.1–1.4, p=.002). By contrast, physically aggressive symptoms tended to decline in frequency with age among boys, though this trend fell just short of significance (POR=.9, 95% CI=.8–1.0, p=.13). Tests for gender differences in 13–15-year-olds suggested that girls with CD tended to show somewhat lower total symptom counts (p=.12) and levels of aggressive symptoms (p=.14) than their male counterparts. Rates of non-aggressive conduct problems were significantly lower among teenage girls than boys (p=.05), but there were no gender differences in levels of status violations (p=.4). Analyses of individual symptoms showed that in terms of aggressive behaviours only fighting – the most commonly endorsed physically aggressive symptom in younger boys –

						showed a significant decline with age (OR=.8, 95% CI=.7–1.0, p=.02); all other aggressive symptoms were reported at essentially similar rates in younger and older boys. Among 13–15-year-olds who met criteria for CD, rates of involvement in aggressive behaviours showed some striking similarities in the symptomatology of girls and boys, with fighting, for example, being reported for 52% of girls and 51% of boys. On other aggressive symptoms there were suggestions of a higher prevalence in boys, though differences only reached conventional levels of statistical significance in relation to cruelty to animals (OR=2.6, 95% CI=1.2–5.8, p=.02). Trends in individual non-aggressive symptoms highlighted a number of interesting features. First, the most marked age-trends arose in symptoms reflecting status violations: levels of truancy (OR=1.3, 95% CI=1.1–1.5, p=.001), staying out late (OR=1.4, 95% CI=1.1–1.6, p=.001) and running away (OR=1.6, 95% CI=1.3–2.2, p < .001) all increased significantly with age. With the exception of fire-setting (which also showed a slight age-related increase: OR=1.1, 95% CI=1.0–1.4, p=.107), rates of all other nonaggressive symptoms – including stealing, breaking in and vandalism – remained essentially stable with age among boys who met criteria for CD. Second, in 13–15-year-olds, rates of individual non-aggressive conduct symptoms and status violations were strikingly similar in boys and girls. The only hints of gender differences arose in relation to breaking in to steal (OR=7.8, 95% CI=1.0–63.3, p=.054) and fire setting (OR=4.6, 95% CI=.9–23.3, p=.062), both of which tended to be more common in boys (pp. 614-615)
Nock, Kazdin, Hiripi, and Kessler	Structured diagnostic interviews	N=3,199; 18-year and older United	Conduct Disorder (CD)	The National Comorbidity Survey	Prevalence rates rise from childhood	The lifetime prevalence of DSM-IV CD is estimated to be 9.5% (12.0% among males and 7.1% among females) (p. 702)

(2006)		States		Replication (NCS-R); The World Health Organization (WHO) Composite International Diagnostic Interview (CIDI)	to adolescence and are higher among males than among females (p. 473)	Lifetime diagnosis of CD is associated with young age, male gender, low educational attainment, being separated or divorced, residing in the Western US, and residing in urban settings (p. 703)
Boylan, Vaillancourt, Boyle, and Szatmari (2007)	Literature review	N=28 articles; children under 18-years-old	Oppositional Defiant Disorder (ODD)	Cross sectional and longitudinal studies in clinics, community, and epidemiologic samples	The disorder appears to be somewhat more prevalent in males than in females (1.4:1) prior to adolescence. This male predominance is not consistently found in samples of adolescents or adults (p. 464)	Gender differences have been inconsistent across smaller studies. In the largest published study (N = 10,000), the prevalence of ODD appears greater in boys (3.2%) than girls (1.8 %) across ages 6–16. This sex difference appears to vary by age, with higher rates in boys prior to adolescence, and no sex differences in adolescence. There may also be peaks in prevalence around age 7 for boys, and a second peak for both sexes at age 14–15 (p. 487)
Nock, Kazdin, Hiripi, and Kessler (2007)	Structured diagnostic interviews	N=3,199; 18-44-years-old United States	Oppositional Defiant Disorder (ODD)	The National Comorbidity Survey Replication (NCS-R); The World Health Organization (WHO) Composite International Diagnostic Interview (CIDI)	The disorder appears to be somewhat more prevalent in males than in females (1.4:1) prior to adolescence. This male predominance is not consistently found in samples of adolescents or adults (p. 464)	Estimated prevalence is not significantly different for males (11.2%) relative to females (9.2%; z=1.41, *ns*)… Median duration of ODD is 6 years and does not vary greatly either by sex or by age at interview. It is consistently higher among men than women in each age group, although none of these differences are statistically significant (p. 705)

Appendix C

BASC-2 Subscale Descriptions

Activities of Daily Living	The skills associated with performing basic everyday tasks in an acceptable and safe manner.
Adaptability	The ability to adapt readily to changes in the environment.
Aggression	The tendency to act in a hostile manner (either verbal or physical) that is threatening to others.
Anxiety	SRP: Feelings of nervousness, worry, and fear; the tendency to be overwhelmed by problems. PRS/TRS: The tendency to be nervous, fearful or worried about real or imagined problems.
Attention Problems	The tendency to be easily distracted and unable to concentrate more than momentarily.
Attitude to School	Feelings of alienation, hostility, and dissatisfaction regarding school.
Attitude to Teachers	Feelings of resentment and dislike of teachers; beliefs that teachers are unfair, uncaring, or overly demanding.
Atypicality	SRP: The tendency toward bizarre thoughts and behaviors considered "odd." PRS/TRS: The tendency to behave in ways that are considered "odd" or commonly associated with psychosis.
Conduct Problems	The tendency to engage in antisocial and rule-breaking behavior, including destroying property.
Depression	SRP: Feelings of unhappiness, sadness, and dejection; a belief that nothing goes right. PRS/TRS: Feelings of unhappiness, sadness, and stress that may result in an inability to carry out everyday activities or may bring on thoughts of suicide.
Functional Communication	The ability to express ideas and communicate in a way others can easily understand.

Hyperactivity	The tendency to be overly active, rush through work or activities, and act without thinking.
Interpersonal Relations	The perception of having good social relationships and friendships with peers.
Leadership	The skills associated with accomplishing academic, social, or community goals, including the ability to work with others.
Learning Problems	The presence of academic difficulties, particularly in understanding or completing schoolwork.
Locus of Control	The belief that rewards and punishments are controlled by external events or other people.
Relations with Parents	A positive regard towards parents and a feeling of being esteemed by them.
Self-Esteem	Feelings of self-esteem, self-respect, and self-acceptance.
Self-Reliance	Confidence in one's ability to solve problems; a belief in one's personal dependability and decisiveness.
Sensation Seeking	The tendency to take risks and to seek excitement.
Sense of Inadequacy	Perceptions of being unsuccessful in school, unable to achieve one's goals, and generally inadequate.
Social Skills	The skills necessary for interacting successfully with peers and adults in home, school, and community settings.
Social Stress	Feelings of stress and tension in personal relationships; a feeling of being excluded from social activities.
Somatization	The tendency to be overly sensitive to, experience, or complain about relatively minor physical problems and discomforts.
Study Skills	The skills that are conducive to strong academic performance, including organizational skills and good study habits.
Withdrawal	The tendency to evade others to avoid social contact.

Appendix D

Individuals with Disabilities Education Improvement Act of 2004 §618

``(c) Technical Assistance.--The Secretary may provide technical assistance to States to ensure compliance with the data collection and reporting requirements under this title.

``(d) Disproportionality.--

``(1) In general.--Each State that receives assistance under this part, and the Secretary of the Interior, shall provide for the collection and examination of data to determine if significant disproportionality based on race and ethnicity is occurring in the State and the local educational agencies of the State with respect to--

``(A) the identification of children as children with disabilities, including the identification of children as children with disabilities in accordance with a particular impairment described in section 602(3);

``(B) the placement in particular educational settings of such children; and

``(C) the incidence, duration, and type of disciplinary actions, including suspensions and expulsions.

``(2) Review and revision of policies, practices, and procedures.--In the case of a determination of significant disproportionality with respect to the identification of children as children with disabilities, or the placement in particular educational settings of such children, in accordance with paragraph (1), the State or the Secretary of the Interior, as the case may be, shall--

``(A) provide for the review and, if appropriate, revision of the policies, procedures, and practices used in such identification or placement to ensure that such policies,

[[Page 118 STAT. 2740]]

procedures, and practices comply with the requirements of this title;

``(B) require any local educational agency identified under paragraph (1) to reserve the maximum amount of funds under section 613(f) to provide comprehensive coordinated early intervening services to serve children in the local educational agency, particularly children in those groups that were

significantly overidentified under paragraph (1); and

``(C) require the local educational agency to publicly report on the revision of policies, practices, and procedures described under subparagraph (A).

Appendix E

Table E1

PRS-A Normality Female (N=513)

			Skewness (Standard Error=.108)		Kurtosis (Standard Error=.215)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	3	24	-.223	-2.06	-.676	-3.144
Activities of Daily Living	1	24	-.437	-4.046	.249	1.158
Aggression	.00	30	2.266	20.981	5.512	25.637
Anxiety	.00	33	.913	8.454	1.240	5.767
Attention Problems	.00	18	.543	5.028	-.438	-2.037
Atypicality	.00	30	2.274	21.056	5.139	23.902
Conduct Problems	.00	42	2.436	22.556	6.399	29.763
Depression	.00	39	1.896	17.556	3.953	18.386
Functional Communication	8	36	-.490	-4.537	-.670	-3.116
Hyperactivity	.00	24	1.789	16.565	3.581	16.656
Leadership	2	30	-.227	-2.102	-.555	-2.581
Social Skills	1	24	-.529	-4.898	-.375	-1.744
Somatization	.00	33	1.872	17.333	3.988	18.549
Withdrawal	.00	21	.994	9.204	.628	2.921

Table E2

PRS-A Normality Male (N=502)

			Skewness (Standard Error=.109)		Kurtosis (Standard Error=.218)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	.00	24	-.268	-2.459	-.474	-2.174
Activities of Daily Living	.00	24	-.257	-2.358	-.293	-1.344
Aggression	.00	28	1.890	17.339	4.170	19.128
Anxiety	.00	29	.622	5.706	.307	1.408
Attention Problems	.00	18	..294	2.697	-.500	-2.294
Atypicality	.00	26	2.118	19.431	5.427	24.894
Conduct Problems	.00	40	2.134	19.578	6.628	30.404
Depression	.00	35	1.670	15.321	3.118	14.303
Functional Communication	4	36	-.461	-4.229	-.426	-1.954
Hyperactivity	.00	24	1.462	13.413	2.532	11.615
Leadership	.00	30	-.121	-1.11	-.497	-2.28
Social Skills	.00	24	-.254	-2.33	-.699	-3.206
Somatization	.00	30	1.930	17.706	5.638	25.863
Withdrawal	.00	24	1.052	9.651	1.021	4.683

Table E3

PRS-C Normality Female (N=501)

			Skewness (Standard Error=.109)		Kurtosis (Standard Error=.218)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	2	24	-.339	-3.11	-.373	-1.711
Activities of Daily Living	4	24	-.385	-3.532	-.306	-1.403
Aggression	.00	20	1.347	12.358	2.12	9.725
Anxiety	.00	34	.515	4.725	.434	1.991
Attention Problems	.00	18	.632	5.798	-.23	-1.193
Atypicality	.00	27	2.178	19.981	6.842	31.385
Conduct Problems	.00	20	1.083	9.936	2.146	9.844
Depression	.00	30	1.735	15.917	4.263	19.555
Functional Communication	4	36	-.695	-6.376	.244	1.119
Hyperactivity	.00	30	1.302	11.945	2.541	11.656
Leadership	.00	24	-.180	-1.651	-.396	-3.193
Social Skills	1	24	-.514	-4.716	-.201	-0.922
Somatization	.00	22	1.064	9.761	1.596	7.321
Withdrawal	.00	25	1.012	9.284	.847	3.885

Table E4

PRS-C Normality Male (N=531)

			Skewness (Standard Error=.106)		Kurtosis (Standard Error=.212)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	1	24	-.053	-0.5	-.665	-3.137
Activities of Daily Living	1	24	.251	2.368	-.172	-0.811
Aggression	.00	23	1.241	11.708	1.609	7.59
Anxiety	.00	38	.756	7.132	1.086	5.123
Attention Problems	.00	16	.014	0.132	-.798	-3.764
Atypicality	.00	25	1.64	15.472	2.56	12.075
Conduct Problems	.00	22	1.294	12.208	2.4	11.321
Depression	.00	34	1.763	16.632	3.562	16.802
Functional Communication	1	36	-.412	-3.887	-.433	-2.042
Hyperactivity	.00	29	1.008	9.509	.934	4.406
Leadership	.00	24	.054	0.509	-.441	-2.08
Social Skills	.00	24	-.126	-1.189	-.657	-3.099
Somatization	.00	24	1.223	11.538	1.913	9.024
Withdrawal	.00	30	1.267	11.953	1.948	9.189

Table E5

PRS-P Normality Female (N=359)

			Skewness (Standard Error=.129)		Kurtosis (Standard Error=.257)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	2	24	-.079	-.612	-.541	-2.105
Activities of Daily Living	2	26	-.413	-3.202	-.044	-0.171
Aggression	.00	33	1.97	15.271	9.363	36.432
Anxiety	.00	35	.94	7.287	2.429	9.451
Attention Problems	.00	17	.156	1.209	-.341	-1.327
Atypicality	.00	28	2.146	16.636	8.121	31.599
Depression	.00	33	1.407	10.907	5.535	21.537
Functional Communication	2	33	-.189	-1.465	-.594	-2.311
Hyperactivity	.00	32	1.001	7.76	1.381	5.374
Social Skills	1	26	-.181	-1.403	-.676	-2.63
Somatization	.00	39	2.142	16.605	14.026	54.576
Withdrawal	.00	23	-..417	-3.233	.132	.514

Table E6

PRS-P Normality Male (N=387)

			Skewness (Standard Error=.124)		Kurtosis (Standard Error=.247)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	.00	24	-.100	-.806	-.459	-1.858
Activities of Daily Living	1	27	-.239	-1.927	-.547	-2.215
Aggression	.00	33	1.845	14.879	6.790	27.49
Anxiety	.00	32	1.050	8.468	1.590	6.437
Attention Problems	.00	17	.237	1.911	-.393	-1.591
Atypicality	.00	30	2.146	17.306	7.963	32.239
Depression	.00	28	.919	7.411	2.143	8.676
Functional Communication	2	33	-.219	-1.766	-.732	-2.964
Hyperactivity	.00	32	1.031	8.315	1.472	5.96
Social Skills	.00	27	-.251	-2.024	-.445	-1.802
Somatization	.00	26	1.257	10.137	2.514	10.178
Withdrawal	.00	27	.714	5.758	.620	2.51

Table E7

SRP-A Normality Female (N=568)

			Skewness (Standard Error=.103)		Kurtosis (Standard Error=.205)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Anxiety Problems	.00	35.00	.604	5.864	-.247	-1.205
Attitude to School	.00	19.00	.927	9	.606	2.956
Attitude to Teachers	.00	25.00	1.366	13.262	2.715	13.244
Attention Problems	.00	21.00	.692	6.718	-.392	-1.912
Atypicality	.00	23.00	1.886	18.311	3.755	18.317
Depression	.00	28.00	1.754	17.029	2.803	13.673
Hyperactivity	.00	20.00	.676	6.563	.429	2.093
Interpersonal Relations	.00	19.00	-1.582	-15.359	2.912	14.205
Locus of Control	.00	22.00	.971	9.427	.523	2.551
Relations with Parents	.00	29.00	-.710	-6.893	.089	.4341
Self-Esteem	.00	20.00	-1.264	-12.272	1.048	5.112
Sensation Seeking	.00	24.00	.279	2.709	-.316	-1.541
Sense of Inadequacy	.00	26.00	1.193	11.583	1.407	6.863
Social Skills	.00	28.00	1.008	9.786	1.120	5.463
Somatization	.00	15.00	1.161	11.272	.656	3.2
Self-Reliance	3	22	-.489	-4.748	-.139	-.678

Table E8
SRP-A Normality Male (N=576)

			Skewness (Standard Error=.102)		Kurtosis (Standard Error=.203)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Anxiety Problems	.00	30.00	.562	5.51	-.296	-1.458
Attitude to School	.00	19.00	.699	6.853	-.102	-.502
Attitude to Teachers	.00	23.00	.714	7	.287	1.414
Attention Problems	.00	21.00	.329	3.225	-.782	-3.852
Atypicality	.00	21.00	1.601	15.696	2.504	12.335
Depression	.00	25.00	1.758	17.235	3.068	15.113
Hyperactivity	.00	19.00	.433	4.245	-.357	9.066
Interpersonal Relations	1.00	19.00	-1.262	-12.373	1.739	8.567
Locus of Control	.00	20.00	.793	7.775	.274	1.35
Relations with Parents	1.00	29.00	-.657	-6.441	.037	.182
Self-Esteem	1.00	20.00	-1.426	-13.98	2.214	10.906
Sensation Seeking	.00	24.00	-.064	-.6275	.053	.261
Sense of Inadequacy	.00	20.00	.751	7.363	.131	.645
Social Skills	.00	20.00	.630	6.176	-.289	-1.424
Somatization	.00	14.00	1.505	14.755	2.012	9.911
Self-Reliance	.00	22	-.472	-4.627	.075	.369

Table E9

SRP-C Normality Female (N=333)

			Skewness (Standard Error=.134)		Kurtosis (Standard Error=.266)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Anxiety Problems	.00	37.00	.735	5.485	.340	1.278
Attitude to School	.00	18.00	1.378	10.284	1.897	7.132
Attitude to Teachers	.00	16.00	1.638	12.224	2.435	9.154
Attention Problems	.00	24.00	.824	6.149	.259	.974
Atypicality	.00	23.00	1.150	8.582	.788	2.962
Depression	.00	30.00	1.916	14.299	3.825	14.38
Hyperactivity	.00	23.00	.771	5.754	.389	1.462
Interpersonal Relations	.00	16.00	-2.235	-16.679	6.121	23.011
Locus of Control	.00	20.00	1.029	7.679	.720	2.707
Relations with Parents	1.00	27.00	-.908	-6.776	.290	1.09
Self-Esteem	1.00	19.00	-1.900	-14.179	4.019	15.109
Sense of Inadequacy	.00	22.00	1.286	9.597	2.031	7.635
Social Skills	.00	21.00	1.091	8.142	1.059	3.981
Self-Reliance	3.00	22.00	-.888	-6.627	.803	3.019

Table E10
SRP-A Normality Male (N=371)

			Skewness (Standard Error=.127)		Kurtosis (Standard Error=.253)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Anxiety Problems	.00	32.00	.410	3.228	-.117	-.462
Attitude to School	.00	18.00	.637	5.016	-.455	-1.798
Attitude to Teachers	.00	19.00	1.452	11.433	2.473	9.775
Attention Problems	.00	21.00	.189	1.488	-.759	-3
Atypicality	.00	24.00	.737	5.803	.132	.522
Depression	.00	30.00	1.339	10.543	1.834	7.249
Hyperactivity	.00	23.00	.506	3.984	.120	.474
Interpersonal Relations	2.00	16.00	-1.506	-11.858	1.967	7.775
Locus of Control	.00	19.00	.555	4.37	-.204	-.806
Relations with Parents	3.00	27.00	-.778	-6.126	.062	.245
Self-Esteem	3.00	19.00	-1.492	-11.748	2.210	8.735
Sense of Inadequacy	.00	19.00	.738	5.811	.971	3.838
Social Skills	.00	20.00	.624	4.913	-.005	-.0198
Self-Reliance	4.00	22.00	-.164	-1.291	-.581	-2.296

Table E11

TRS-A Normality Female (N=272)

			Skewness (Standard Error=.148)		Kurtosis (Standard Error=.294)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	3.00	24.00	-.549	-3.709	-.454	-1.544
Aggression	.00	20.00	2.492	16.838	7.507	25.534
Anxiety	.00	20.00	1.426	9.635	2.767	9.412
Problems Attention Problems	.00	20.00	1.034	6.986	.684	2.327
Atypicality	.00	17.00	3.278	22.149	15.214	5.748
Conduct Problems	.00	19.00	2.209	14.926	4.764	16.204
Depression	.00	22.00	2.150	14.527	6.549	22.276
Functional Communication	1.00	24.00	-.807	-5.453	.091	.32
Hyperactivity	.00	26.00	1.851	12.507	3.792	12.898
Leadership	1.00	18.00	-.138	-.932	-.766	-2.605
Learning Problems	.00	20.00	1.869	12.628	3.761	12.793
Social Skills	1.00	24.00	-.266	-1.797	-.760	-2.585
Somatization	.00	18.00	2.671	18.047	9.281	31.568
Study Skills	2.00	27.00	-.542	-3.662	-.662	-2.252
Withdrawal	.00	17.00	1.536	10.378	2.570	8.741

Table E12

TRS-A Normality Male (N=272)

			Skewness (Standard Error=.148)		Kurtosis (Standard Error=.294)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	1.00	24.00	-.379	-2.561	-.584	-1.986
Aggression	.00	28.00	1.706	11.527	2.974	10.116
Anxiety Problems	.00	18.00	1.478	9.986	2.320	7.891
Attention Problems	.00	21.00	.332	2.243	-.896	-3.048
Atypicality	.00	18.00	1.753	11.845	3.112	10.585
Conduct Problems	.00	28.00	1.573	10.628	2.390	8.129
Depression	.00	20.00	1.620	10.945	2.446	8.32
Functional Communication	.00	24.00	-.449	-3.034	-.631	-2.146
Hyperactivity	.00	31.00	1.227	8.291	1.132	3.85
Leadership	.00	18.00	.287	1.939	-.825	-2.806
Learning Problems	.00	22.00	1.074	1.257	.738	2.51
Social Skills	.00	24.00	.006	.041	-.900	-3.061
Somatization	.00	15.00	2.072	14	4.580	15.578
Study Skills	.00	27.00	.010	.068	-1.198	-4.075
Withdrawal	.00	20.00	1.146	7.743	1.105	3.759

Table E13

TRS-C Normality Female (N=411)

			Skewness (Standard Error=.120)		Kurtosis (Standard Error=.240)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	2.00	24.00	-.640	-5.333	-.090	-.375
Aggression	.00	18.00	2.461	20.508	6.245	26.021
Anxiety Problems	.00	18.00	1.297	10.808	2.285	9.521
Attention Problems	.00	20.00	.757	6.308	-.292	-1.217
Atypicality	.00	17.00	2.506	20.883	7.935	33.063
Conduct Problems	.00	22.00	2.073	17.275	5.478	22.825
Depression	.00	21.00	2.113	17.608	5.851	24.379
Functional Communication	3.00	30.00	-.709	-5.908	-.039	.163
Hyperactivity	.00	30.00	1.652	13.767	3.130	13.042
Leadership	.00	18.00	-.111	-.925	-.733	-3.054
Learning Problems	.00	21.00	1.563	13.025	2.310	9.625
Social Skills	.00	24.00	-.368	-3.067	-.415	-1.729
Somatization	.00	21.00	2.145	17.875	5.524	4.345
Study Skills	.00	21.00	-.558	-4.65	-.409	-1.704
Withdrawal	.00	20.00	1.763	14.692	3.661	15.254

Table E14
TRS-C Normality Male (N=378)

			Skewness (Standard Error=.125)		Kurtosis (Standard Error=.250)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	3.00	24.00	-.151	-1.208	-.821	-3.284
Aggression	.00	29.00	1.782	14.256	3.773	15.092
Anxiety Problems	.00	18.00	1.429	11.432	2.502	10.008
Attention Problems	.00	21.00	.162	1.296	-.992	-3.968
Atypicality	.00	21.00	1.812	14.496	3.520	14.08
Conduct Problems	.00	19.00	1.289	10.312	1.457	5.828
Depression	.00	18.00	1.587	12.696	2.375	9.5
Functional Communication	1.00	30.00	-.515	-4.12	-.411	-1.644
Hyperactivity	.00	31.00	.840	6.72	.118	.472
Leadership	.00	18.00	.172	.727	-.688	-2.752
Learning Problems	.00	20.00	1.002	8.016	.275	1.1
Social Skills	.00	24.00	.055	.44	-.672	-2.688
Somatization	.00	19.00	1.938	15.504	4.661	18.644
Study Skills	.00	21.00	-.180	-1.44	-.889	-3.556
Withdrawal	.00	23.00	1.396	11.168	1.991	7.964

Table E15

TRS-P Normality Female (N=241)

			Skewness (Standard Error=.157)		Kurtosis (Standard Error=.312)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	.00	21.00	-.295	-1.879	-.846	-2.729
Aggression	.00	26.00	1.923	12.248	4.639	14.869
Anxiety Problems	.00	15.00	.843	5.369	.393	1.26
Attention Problems	.00	17.00	.413	2.631	-.782	-2.506
Atypicality	.00	22.00	2.395	15.255	7.945	25.465
Depression	.00	20.00	1.185	7.548	2.079	6.663
Functional Communication	1.00	27.00	-.113	-.72	-.730	-2.34
Hyperactivity	.00	23.00	1.389	8.847	2.200	7.051
Social Skills	.00	18.00	-.031	-.197	-.749	-2.401
Somatization	.00	22.00	1.118	7.121	2.998	9.609
Withdrawal	.00	16.00	.883	5.624	.470	1.506

Table E16
TRS-P Normality Male (N=210)

			Skewness (Standard Error=.168)		Kurtosis (Standard Error=.334)	
Subscale	Minimum	Maximum	Skewness	Normality z-score	Kurtosis	Normality z-score
Adaptability	1.00	21.00	-.130	-.774	-.879	-2.632
Aggression	.00	31.00	1.591	9.47	2.686	8.042
Anxiety Problems	.00	18.00	1.219	7.256	1.964	5.88
Attention Problems	.00	18.00	.201	1.196	-.724	-2.168
Atypicality	.00	17.00	1.569	9.339	2.018	6.042
Depression	.00	20.00	1.076	6.405	1.233	3.692
Functional Communication	.00	27.00	-.057	-.339	-.724	-2.168
Hyperactivity	.00	26.00	1.198	7.131	1.455	4.356
Social Skills	.00	18.00	.367	2.185	-.707	-2.117
Somatization	.00	11.00	.539	3.208	-.632	-1.892
Withdrawal	.00	23.00	1.650	9.821	3.969	11.883

Appendix F

Hypotheses, Variables, Analysis Matrix

Hypotheses	Independent Variables	Dependent Variables	Analyses
H1: The genders will be similarly rated on *the Behavior Assessment System for Children, second edition Parent Rating Scale* (BASC-2 PRS).	Gender.	BASC-2 PRS: Subscale Raw Scores.	Weber and Popova Tests of Equivalence; Correlation Coefficient; Coefficient of Determination; Mean; Standard Deviation.
H2: The genders will be similarly rated on *the Behavior Assessment System for Children, second edition Teacher Rating Scale* (BASC-2 TRS).	Gender.	BASC-2 TRS: Subscale Raw Scores	Tests of Equivalence; Effect Sizes.
H3: The genders will be similarly rated on *the Behavior Assessment System for Children, second edition Self Report of Personality* (BASC-2 SRP).	Gender.	BASC-2 SRP: Subscale Raw Scores	Weber and Popova Tests of Equivalence; Correlation Coefficient; Coefficient of Determination; Mean; Standard Deviation.