EXAMINING THE VALIDITY OF THE AUTISM SPECTRUM RATING SCALES WITH A PRESCHOOL POPULATION REFERRED FOR SPECIAL EDUCATION EVALUATION

by

Sean Cunningham

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Educational Psychology

The University of Utah

August 2012

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Sean Cunningham**

has been approved by the following supervisory committee members:

Elaine Clark	, Chair	**5-21-2012** Date Approved
Dan Woltz	, Member	**5-21-2012** Date Approved
Janiece Pompa	, Member	**5-21-2012** Date Approved
Lora Tuesday-Heathfield	, Member	**5-21-2012** Date Approved
Peter Nicholas	, Member	**5-21-2-12** Date Approved

and by **Elaine Clark**, Chair of the Department of **Educational Psychology**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

Early identification and intervention of an Autism Spectrum Disorder (ASD) can have beneficial effects that extend into later life. However, currently used instruments have difficulties detecting children who may have an ASD. The current study investigated the utility of a newly published measure, Autism Spectrum Rating Scales (ASRS). Participants included 67 children ages 2 to 5 years old, referred for possible special education services. Participants were divided into two groups: those with an ASD (n = 37) and others suspected of having a general developmental disability (DD) (n = 30). Participants were assessed using the ASRS to examine the instrument's ability to classify them as having an ASD or a general DD. Additional testing examined the effects various levels of intellectual, adaptive, and language skills have on the ability of the ASRS to classify children. Classification ability and error rates of the ASRS were also examined with regard to base rates and error acceptability by context. Results indicate that with a recommended cut score of 70, the Parent ASRS had an overall hit rate of 64%. The Parent ASRS had a Type I error rate (i.e., false positive) of 16% and a Type II error rate (i.e., false negative) of 19%. For the Teacher ASRS, the hit rate was 62%. The Teacher ASRS had a Type I error rate of 15% and a Type II error rate of 23%. Sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio of the ASRS were also examined to gain insight into the measure's utility. ROC Curve analysis determined the area under the curve (AUC) for the ASRS, the most optimal point for sensitivity and specificity. It was concluded that

across all ASRS forms (e.g., Parent, Teacher), the general ability of the ASRS to classify and discriminate between children with potential ASDs or general DDs referred for possible special education services were similar.

This work is dedicated to my beloved mother, Linda Cunningham, who is a model of strength and an inspiration for everything I do in life. I appreciate all you have done and will do for me. My words and actions will never be able to fully express my gratitude for the life you have created for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I wish to thank the families who dedicated their time and resources to this project including my own family who stood behind me during the course of this project to make it a success. Additionally, I would like to acknowledge the guidance from my dissertation committee at the University of Utah, Elaine Clark, Ph.D., Lora Tuesday Heathfield, Ph.D., Peter Nicholas, Ph.D., Janiece Pompa, Ph.D., and Dan Woltz, Ph.D. My appreciation also goes to those individuals in the Davis School District who have made this project a possibility, including Heidi Block, Ph.D., Benjamin Belnap, Ph.D. and Jennifer Cardinal, Ph.D. An extra special thanks to Whitney Teter for being such an outstanding source of support throughout the course of this project.

CHAPTER I

INTRODUCTION

Overview of Autism Spectrum Disorders

Diagnostic Criteria

Autism Spectrum Disorder (ASD) is a term that describes varying neurodevelopmental conditions in which an individual demonstrates impairments in communication and social interactions, and may display restricted and/or stereotyped behaviors (Johnson, Myers, & Council on Children with Disabilities, 2007). These impairments typically begin to emerge in infancy and toddler years (Ozonoff, Goodlin-Jones, & Soloman, 2005). ASDs are part of a broader classification of Pervasive Developmental Disorders that include Autistic Disorder, Asperger's Disorder, and Pervasive Developmental Disorder-Not Otherwise Specified (APA, 2000; Goldstein, Naglieri, & Ozonoff, 2009). (For full DSM-IV-R criteria on ASDs see Appendix A). Each of these diagnoses is considered to be part of an autism "spectrum" and the symptom severity and behavioral presentation can differ significantly with each individual person (National Research Council, 2001).

Etiology

Autism is generally considered a biologically-based neurodevelopmental disorder with a genetic loading (Autism Genome Project Consortium 2007; Losh, Sullivan, Trembath,

& Piven, 2008). Twin studies indicate upwards of 9 times greater presence of autism in monozygotic twins than dizygotic twins (Bailey et al., 1995). In spite of twin and family studies indicating a genetic link to autism, the specific genetic origin remains elusive; upwards of 90% of ASDs are considered idiopathic (Bertone, Hanck, Kogan, Chaudhuri, & Cornish, 2010). Genetic models typically view the etiology of ASDs as based on a polygenic framework where genetic and environmental factors interact, manifesting to a specific phenotype. Genetic investigations have found a link between ASDs and genetic variations on the chromosome 5p14.1 (Wang et al., 2009). Environmental risk factors have also been associated with autism. A recent meta-analysis by Gardner, Spiegelman, and Buka (2009) indicated that the following environmental conditions are risk factors for the later development of an ASD: maternal and paternal age, birth order, maternal birth abroad, gestational bleeding, gestational diabetes, and maternal medication use.

Base Rates

It is important to acknowledge that base rates of ASDs may differ as a function of environment. More specialized environments providing care for children with ASDs, such as psychologists' offices and special education programming within the public schools, may see higher base rates. Within well-child visits to pediatricians, base rates are more likely reflective of those found in the general population.

Base rates in global population. Estimates of base rates in the global population have been influenced by diagnostic criteria. When considering a more general definition that includes all the pervasive developmental disorder diagnoses of the autism spectrum, rates have been found to be 37 per 10,000 (Fombonne, 2005). When using more restrictive

diagnostic criteria, by considering only the population of children meeting criteria for autism, rates were found to be lower, at 13 per 10,000.

Base rates in the United States population. Within the United States, the most commonly reported prevalence rate of autism spectrum disorders, based on a study done at the Centers for Disease Control and Prevention (CDC), is 6.7 per 1,000 children or approximately 1 in 150 (Rice, 2007). The most recent study published by the CDC, however, indicates that the current rate of autism in the United States is 11.3 per 1,000 children, or approximately 1 in 88 (CDC, 2012). These were retrospective studies, where the children were identified from health and educational records.

Over the past several years, there has been an increasing rate of children diagnosed with ASDs. Data taken from the CDC in 2006 showed a 57% increase in children identified with possible ASDs when compared to data collected in 2002 (Rice, 2009). The most current data released by the CDC in 2012 indicates a 23% increase over the 2006 data and a 78% increase when compared to the 2002 data (CDC, 2012). Research indicates that the number of preschool children with ASDs is rapidly increasing, as well (Chakrabarti & Frombone, 2005; Leonard, Dixon, & Whitehouse, 2010).

Base rates of ASDs in the special education population. Under the Individuals with Disabilities Education Act (IDEA), Part B, 5.8% of the preschool-age population had disabilities and was served through Individualized Education Plans (U.S. Department of Education, 2010). Of these children, 6.8% were identified as having autism, 46.4% had speech or language impairment, and 27.8% had a developmental delay. These numbers indicate that within the special education setting, the base rates of ASDs are higher than the general population. Data from the U.S. Department of Education (2010) show that from 1997

to 2007 there was an increase of over 600% in the number of children served under the special education classification of autism.

Purpose of the Study

As indicated, within the general and special education populations there has been an increase in the identification of children with ASDs. There have been several explanations offered as to the reason for this increase, including a general increase in public and professional awareness of the condition and an increase in the use of instruments used in the identification of ASDs. Instruments used in the identification of ASDs, however, have been criticized due to their lack of sound psychometric properties, and there is concern that their use may lead to the misclassification and over-identification of children on the autism spectrum. This can be further complicated by the fact that at younger ages there is considerable overlap between the symptom presentation of ASDs and other developmental disabilities, such as language delays. Thus, there is a current need to research instruments designed to assess ASDs that can give a clear and accurate diagnostic picture, one that allows for appropriate classification with minimal error. The current study examined the validity of a newly published measure, the Autism Spectrum Rating Scales (ASRS) that is designed to assess preschool children who may be on the autism spectrum, to determine how well the measure accurately classifies preschool children with potential ASDs. The study also set out to determine how well the ASRS discriminates between preschool children who may have an ASD and those with other developmental disabilities.

Current Issues in Autism

Age of Identification

The initial age of recognition of the symptoms of an ASD by parents has been demonstrated to be significantly younger than the initial age of diagnosis of ASDs (Chawarska et al., 2007). Research has shown that parents first notice symptoms of ASD around 14-19 months, which may include: delays in speech and language, abnormal social responsiveness, medical problems, and difficulties with sleeping, eating, and attention.

Although studies are finding that parents notice symptoms of ASDs in the early months of their child's life, they may delay discussing their concerns with their pediatrician for several months (Johnson et al., 2007). Some studies indicate that parents may not seek professional advice until the child is 21 to 25 months old (Gupta et al., 2007). The average age of initial diagnosis of an ASD has been cited at 61 months, or 5 years 1 month (Wiggins, Baio, & Rice, 2006). Thus, the diagnosis of an ASD often occurs significantly after initial parental concerns have arisen.

Early Identification and Intervention

The benefits of early identification of children who may have an ASD have been well documented across several research studies (Dover & LeCouteur, 2007; Goldstein, Naglieri, & Ozonoff, 2009; Rice, 2007; Rogers, 1996, 1998; Smith, Groen, & Wynn, 2000; Turner, Stone, Pozdol, & Coonrod, 2006). Early identification of ASDs is crucial because, as these studies have found, early intervention is one of the best predictors of long-term positive outcomes and result in many benefits, including: education and support to families, services

to address health and behavioral needs, as well as a better prognosis with regard to overall global development. Benefits of early intervention also work to prevent comorbid behavioral difficulties (Dover & LeCouteur, 2007).

Benefits in terms of health. Early identification and subsequent interventions have been demonstrated to improve several domains of a child's development. Ben, Itzchak, and Zachor (2011) described improvements in the areas of verbal abilities, adaptive skills and cognition with early identification and intervention with ASDs. Furthermore, there was a reduction in the presence of behaviors related to the autism spectrum. Other studies have also illustrated the benefits of early intervention on symptoms related to autism spectrum disorder (Granpeesheh, Dixon, Tarbox, Kaplan, & Wilke, 2009; Rickards, Walstab, Wright-Rossi, Simpson, & Reddihough, 2009).

Benefits in terms of cost. There are potential benefits to society for early identification in terms of reduced costs. Ganz (2007) estimates that the lifetime cost for an individual with an ASD is around $3.2 million; however, with early intervention, he claims that this cost can be reduced by $1.4 million. In the public schools alone, effective early intervention with ASDs has been found to generate savings of $187,000 to $203,000 per child (Jacobson, Mulick, & Green, 1998).

Difficulties in the Identification and Intervention of ASDs

The inability of behavioral measures to accurately identify children with ASDs has been cited as the most common reason for delayed diagnosis (Ward & Gilmore, 2010). As discussed above, effective identification is crucial because early identification and intervention of ASDs result in a more positive prognosis. Sound predictive validity has been

demonstrated in the identification of children with a potential ASD as early as age 2, with that prediction still holding at age 9 (Lord et al., 2006). This phenomenon highlights the importance of finding an effective instrument that is able to identify children with ASD early in life.

Difficulties in the identification of ASDs are related to different factors across several contexts of the child's life. From early on, parents are late to report symptoms to a healthcare professional. Although parents may recognize the presentation of symptoms around the age of 1 year, it may not be until around age 2 that those concerns are sought out for professional advice (Gupta et al., 2007). Additionally, funding as well as actual identification methods have been problematic in facilitating the early identification of children who may have an ASD.

Difficulties with Funding

Federal spending on ASDs has increased in recent years, primarily within agencies outside of the U.S. Department of Education. The National Institutes of Health (NIH) has had an increase in spending related to ASDs from $22 million to $74 million from 1997-2002, and the Centers for Disease Control and Prevention has had an increase in spending of $9 million (Verstag, 2004). Despite increases in federal spending over the past several years, schools have not necessarily benefited from these monetary increases. As a result, there have been problems in the school systems with early identification of children with ASDs.

Difficulties with Resources

School personnel that have some level of training in the assessment and identification of ASDs include school psychologists, special education teachers, and speech and language pathologists. However, limited resources in the schools have compromised the ability of school personnel to appropriately assess and identify children with ASDs (Sikora, Hall, Hartley, Gerrard-Morris, & Cagle, 2008). Schools often fail to spend enough money on materials and/or training to assist in the identification of ASDs (Sikora et al., 2008). As a result, Sikora and colleagues (2008) indicate that in place of more comprehensive assessment methods for ASDs, schools are increasingly utilizing parent behavior checklists, which require little time and money to administer, as a primary means to assess and identify children with ASDs. This practice is particularly problematic, given findings that many parent behavior checklists attempting to identify ASDs have compromised psychometric properties, including weak predictive validity (Campbell, 2005; Lecavalier, 2005; National Research Council, 2001; Perry, Condillac, Freeman, Dunn-Geier, & Belair, 2005). Consequently, these checklists have led to unreliable identification and misidentification of children with ASDs (Campbell, 2005).

Schools as a Context in the Identification of ASDs

Federal law mandates that school children with an ASD be identified to determine if special education services may be warranted under the protection of the Individuals with Disabilities Education Improvement Act of 2004 (IDEA; U.S. Dept. of Education, 2004). IDEA operates to assure a Free and Appropriate Public Education (FAPE) for all individuals with disabilities. Part C of IDEA covers early intervention services for children ages birth to

2, whereas Part B covers children's services from ages 3 – 21. Children evaluated under the guidelines of IDEA Parts B and C are typically referred by an educational professional or parent.

Increase of ASDs in Schools

As in the general population, there have been an increased number of children with ASDs in the public schools (Shattuck, 2002). In fact, the rate of children served under the special education classification of Autism in the public schools has continued to increase since the introduction of Autistic Disorder as a separate diagnostic category with the publication of the DSM-III-R (American Psychiatric Association, 1987). In examining a decade of data, the U.S. Department of Education (2010) reported that during the 1997-1998 school year, 42 children were being served under the classification of Autism compared to the 2007-2008 school year, during which 296 students were served under the same category. This is an increase of over 600%. Perhaps more alarming is that the population of children served under the classification of Autism in the public schools is most likely underestimated because children with multiple disabilities (e.g., Autism and Intellectual Disability) are only classified by their primary disability (Newschaffer, Falb, & James, 2005). Therefore, children meeting criteria for multiple special education classifications (e.g., autism, intellectual disability, speech and language impaired) may not always be classified, counted, or served under the Autism classification.

Identification of ASDs in the Schools

School psychologists generally sit on a multidisciplinary team that plays a role in the assessment and identification of children with ASDs. Their role on the team may be to conduct intellectual, behavioral, and/or adaptive behavior assessments. There are several circumstances, including time and training limitations, which often limit a school psychologist's ability to perform comprehensive assessments that follow best practice guidelines. As a result, in lieu of more comprehensive evaluations following best practice guidelines, behavioral checklists have often been utilized as a primary means to identify children with ASDs (Sikora et al., 2008). This, in and of itself, is problematic, but when the measures used to assess for ASDs have psychometric weaknesses, the practice becomes even more problematic. Thus, there is a current need for a psychometrically sound measure that increases the chances of accurately identifying ASDs and ensures that those with other developmental disabilities are not falsely identified (South et al., 2002).

ASD Screening Measures in the Schools

Of the instruments that are being used to differentiate between children on the autism spectrum and those children with a general developmental disability (DD), the most widely used measures in the schools have little research published about their utility to identify children with an ASD over children with other DDs. Allen (2008) found that the most commonly used behavioral measures in the identification and classification of children with ASDs in the schools included the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1988), Gilliam Autism Rating Scale (GARS; Gilliam, 1995), and Gilliam Asperger's Disorder Scale (GADS; Gilliam, 2001), respectively.

Childhood Autism Rating Scale (CARS)

The CARS has been criticized as being dated, as it is based on now outdated diagnostic criteria for ASDs (National Research Council, 2001). Further, the CARS may not be helpful in early identification because it was not designed for use with preschoolers (Vig & Jedrysek, 1999). Additionally, the CARS has not been shown to reliably differentiate between autism and other developmental conditions (Perry et al., 2005). The CARS was revised in 2010 (CARS-2; Schopler, Van Bourgondien, Wellman, & Love, 2010), but to date there are no studies that examine the utility of the revised version.

Gilliam Autism Rating Scale (GARS)

Another commonly used behavioral measure in the schools, the GARS, has been criticized for its psychometric properties and tendency to underestimate children with ASDs with large percentages of false negatives (Lecavalier, 2005; Mazefsky & Oswald, 2006). Additionally, the GARS has difficulty differentiating between ASDs and other developmental disabilities (e.g., communication disorders) (Mazefsky & Oswald, 2006). Researchers have expressed caution in using the GARS in the identification process of children with ASDs (South et al., 2002). Of note is the fact that the GARS has been revised (i.e., GARS-2) (Gilliam, 2006); however this investigator could not find any empirical studies that have examined the psychometric properties of the measure, aside from those published by the developer in the initial instrument's development and standardization.

Gilliam Asperger's Disorder Scale (GADS)

The GADS, another checklists frequently used in the schools to identify children with potential ASDs, has also been criticized for its compromised psychometric properties (Campbell, 2005). Moreover, this measure is used only with higher-functioning individuals, such as those with Asperger's Disorder or high-functioning autism. Later identification of ASDs is less beneficial, because, as discussed previously, it is important to have behavioral measures that accurately identify children with ASDs early in life in order to ensure that these children have the best chance of receiving early, and appropriate intervention (Jacobson et al., 1998; Rice, 2007; Rogers, 1996; Smith et al., 2000).

The Autism Spectrum Rating Scales (ASRS)

The Autism Spectrum Rating Scales (ASRS; Goldstein & Naglieri, 2010) is a newly developed behavior measure that attempts to identify children with ASDs and distinguish children with ASDs from populations of children with other disabilities. Because it is new, it remains in limited use and has yet to be fully tested beyond the initial work done by its developers. The ASRS has potential to be used to identify children with ASDs in several different contexts. The ASRS relies on parent and teacher ratings of students ages 2 through 18 years old. There are two forms for each age group, one for ages 2 through 5 and another for ages 6 through 18. For the younger age group, there are Parent and Teacher/Childcare Provider Forms. Each form asks the rater to rate the students' behavior on a Likert-style scale. The Parent and Teacher/Childcare Provider form have the same 70 items. Table 1.1 outlines the psychometric properties of the top three most commonly used instruments used in the identification of children who may have an ASD as used in the schools, along with the

Table 1.1: *Comparison of Commonly Used Measures in the School Setting with the ASRS*

Measure	Sensitivity	Specificity	Cronbach's alpha
ASRS[a]	89.8 - 95.4	88.6 - 94.5	.97
CARS[b]	.88	.86	.94
CARS-2[c]	.79-.81	.58-.87	.96
GADS[d]	.75-.96	.60-.80	.87
GARS[e]	*	*	.96
GARS-2[f]	.84-1.00	.84-.87	.94

Note: ASRS= Autism Spectrum Rating Scales; CARS = Childhood Autism Rating Scale; CARS-2; Childhood Autism Rating Scale – Second Edition; GADS = Gilliam Asperger's Disorder Scale; GARS = Gilliam Autism Rating Scale; GARS-2 = Gilliam Autism Rating Scale – Second Edition. Reliability coefficients reflect internal consistency and interrater reliability. Validity coefficients reflect concurrent validity. [a] Goldstein & Naglieri, 2010; [b] Schopler, et al., 1980; [c] Schopler, et al., 2010 [d] Gilliam, 2003; [e] Gilliam, 1995; [f] Gilliam, 2006
*Not reported

ASRS, the focus of the current research study. It should be noted that the populations in the standardization samples for each measure vary considerably and ranges in sensitivity and specificity coefficients reflect the range that stems from those differences. Some of the samples include children with a DD while others do not. Moreover, given that the nature of other forms of validity evidence is so varied for these instruments, this was not included in the table. As can be seen in Table 1.1, the ASRS evidence from the standardization sample compares favorably with the other existing instruments

The ASRS may have benefits over other behavior checklists that seek to identify children with potential ASDs. One component of the ASRS that is lacking in other currently published scales is a set of individual scales, including a total aggregate scale, two subscales (e.g., Social/Communication, Unusual Behavior), a scale that matches DSM-IV-TR criteria

and individual treatment scales. The scale structure of the ASRS allows it to be used for initial screenings, progress monitoring, and more thorough evaluations.

Two forms of the ASRS are available, the Full-Length Form and the Short Form. The Full-Length ASRS contains an overall Total Score, which is a T-score. This is comprised of two subscale scores: Unusual Behaviors and Social/Communication. These subscale scores are comprised of Treatment Scales, which include the following: Peer Socialization, Adult Socialization, Social/Emotional Reciprocity, Atypical Language, Stereotypy, Behavioral Rigidity, Sensory Sensitivity, and Attention/Self-Regulation. There is also a separate scale that includes the diagnostic symptoms of ASDs from the DSM-IV-TR. The Short Form of the ASRS is comprised of the 15 items taken from the Full-Length Form that were best able to distinguish children with ASDs from typically developing children. See Figure 1.1 for an overview of the scale structure of the ASRS.

The ASRS was standardized using a wide range of children aimed to match the general population of the United States. The normative sample of the ASRS included 2,560 children highly reflective of the most recent U.S. Census (Goldstein & Naglieri, 2010). Of these children, 640 were utilized for the ASRS form for ages 2 through 5. Table 2.1 outlines the normative data for the ASRS (ages 2-5).

Differentiating ASDs from Other DDs in the Schools

Most developmental disabilities, including ASDs, have similarities in their overt symptoms and behavioral presentation with other developmental disabilities that may present in the school context (Baron-Cohen, Cox, Baird, Swettenham, & Nightingale, 1996; Charman et al., 1998; Lord, 1995; Ventola et al., 2007). This symptom overlap complicates

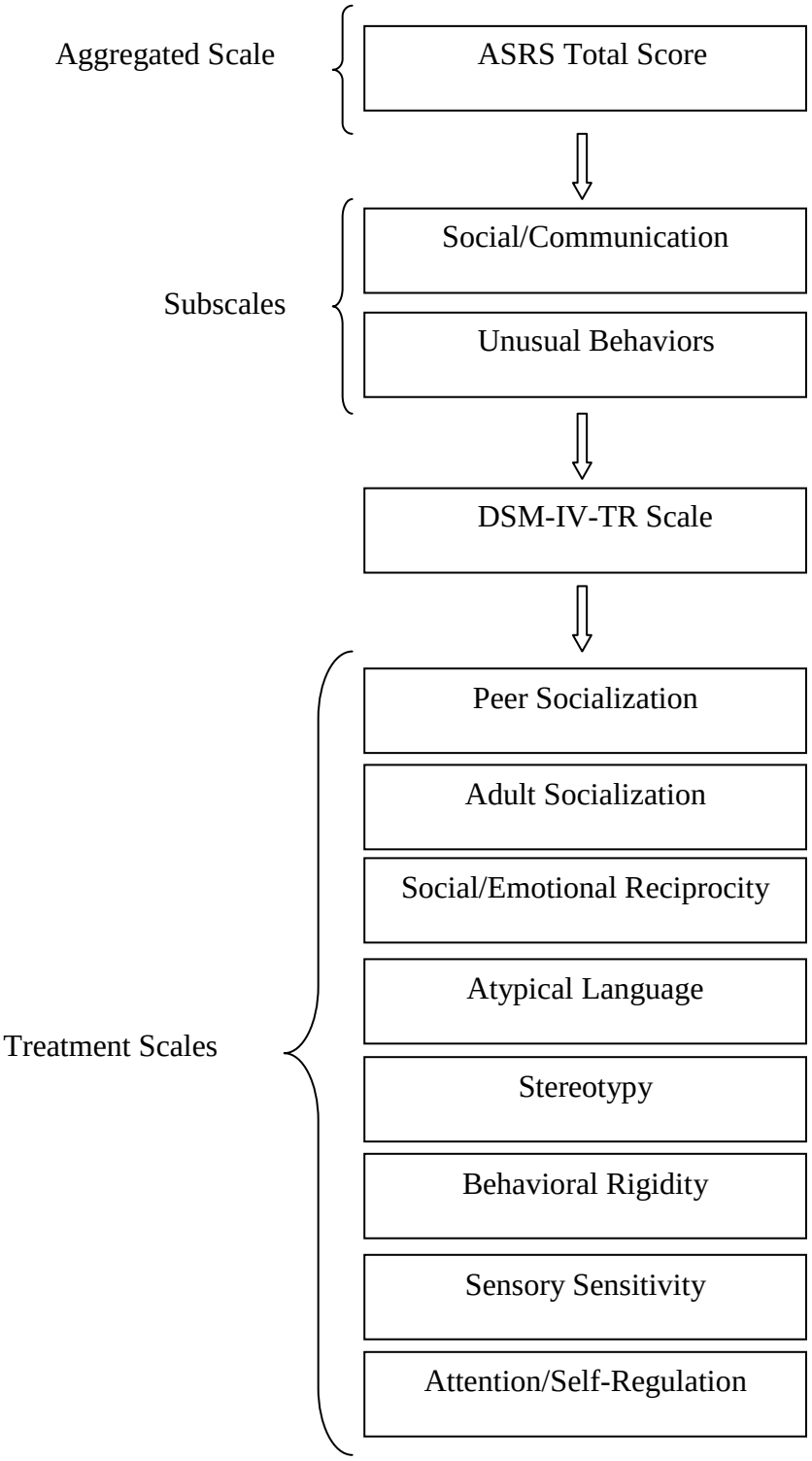

Figure 1.1 *Autism Spectrum Rating Scales Scale (ASRS, 2-5) Structure*

Table 2.1: *Demographic Normative Data for the ASRS (2-5)*

		Parent (*N* = 320)	Teacher (*N* = 320)
Age			
	2	80	80
	3	80	80
	4	80	80
	5	80	80
Gender			
	Male	160	160
	Female	160	160
Ethnicity			
	Asian	10	8
	African/ American	53	48
	Hispanic	58	48
	White	184	199
	Other	15	17

Note: ASRS = Autism Spectrum Rating Scales

Source: Goldstein & Naglieri, 2010

the identification of children in the schools who may be on the autism spectrum. Examples of developmental disabilities that have similarities in their behavioral presentation with ASDs include intellectual disability (ID) and speech and language impairment (SLI). IDEA definitions for these disabilities can be found in Appendix B.

As a result of the similarities in behavioral presentation, instruments that assess for possible ASDs may have difficulty differentiating ASDs from other disabilities. Moreover, the similarities in behavioral presentation may also lead to more error in the detection of

children with ASDs (e.g., Type I or false positive & Type II or false negative). In particular, differentiating ASDs from other forms of developmental disabilities has been demonstrated to be particularly problematic in populations of children under 6 years of age (Vig & Jedrysek, 1999). Additionally, there is currently a paucity of autism screening instruments used to differentiate young children with ASDs from children with other developmental disabilities (Ward, & Gilmore, 2010).

Of the research that is published, results indicate that many currently used measures to identify children with ASDs do not accurately differentiate between children on the autism spectrum and those with other forms of a DD, particularly in younger children (Chawarska, Klin, Paul, & Volkmar, 2007; Gray, Tonge, & Sweeney, 2008; Lord, Storoschuk, Rutter, & Pickles, 1993; Saemundsen, Magnusson, Smari, & Sigurdardottir, 2003). In a review of the literature, a few published studies have explored some measures with the preschool population. Ventola et al. (2007) found that items of the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) and CARS were better able to differentiate between groups of children with an ASD and a general DD than did the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999). The ability of the ADOS and CARS to better differentiate between groups of children with an ASD and a general DD was attributed to the fact that items on both measures addressed more socialization and social interaction factors. The Social Communication Questionnaire (SCQ; Berument, Rutter, Lord, Pickles & Bailey, 1999) has been used to screen for children with possible ASDs. However, when examining its utility, Allen, Silove, Williams, and Hutchins (2007) found that the SCQ results in a high number of false positives with preschool-aged children.

Effective Identification Instruments

The ADOS has been used in research with preschoolers and has been demonstrated to be successful at providing information that effectively aids in the early classification of ASDs, including aiding in differential diagnosis. In the current literature on ASDs, several research studies have examined the ADOS and found that the ADOS has sound psychometric properties in identifying children with ASDs (Gray, Tonge, & Sweeney, 2008; Le Couteur, Haden, Hammal, & McConachie, 2008). Moreover, the ADOS has been used in research to classify preschool children with ASDs, and has been found to be effective in differential diagnosis with other disabilities in this age group (Gray, Tonge, & Sweeney, 2008). The ADOS has also been utilized in research for identifying preschool children with ASDs in educational settings (Lee, David, Rusyniak, Landa, & Newschaffer, 2007).

Unique Behavioral Presentation

Although, as mentioned previously, there is overlap in the behavioral presentation of children with ASDs and other disabilities, there are also unique behavioral differences between these groups of children. Compared to children with other disabilities, children with ASDs display greater impairment in social interaction (e.g., joint attention, quality of social overtures, shared enjoyment) (Lord, 1995; Lord, & Pickles, 1996; Noterdaeme, Sitter, Mildenberger, & Amorosa, 2000; Rogers, Hepburn, Stackhouse, & Wehner, 2003), communication (e.g., echolalia) (Landry, & Loveland, 1988; Lord, Storoschuk, Rutter, & Pickles, 1993; Trillingsgaard, Sorensen, Nemec, & Jorgensen, 2005), and play (e.g., pretend, joint attention) (Charman et al., 1997; Cox et al., 1999; Lord, Rutter, & Le Couteur, 1994). Beyond these specific areas, children with ASDs tend to score lower across multiple

developmental domains, including language skills, visual perception, fine motor skills, and adaptive behavior (Ventola et al., 2007). Finding a specific behavioral measure for autism that accurately differentiates between children with ASDs and children with other developmental delays, however, has been a challenge.

Best Practice Strategies

Given suggested best practice standards and research on instruments that assess ASDs, professionals have utilized a diverse group of strategies to differentiate between children with ASDs and other developmental disabilities. However, the methods that aid in the differential diagnosis of children with ASDs are not standardized. Researchers have suggested a comprehensive assessment approach when attempting differential diagnosis with children having ASDs. Assessment of language and social skill impairments, as well as direct observations, behavior checklists, interviews with parents and caregivers, and other measures should be utilized in order to aid in decisions regarding differential diagnosis (Matson, 2007). Other researchers have suggested examining variables that are maintaining the behaviors of interest, examining skill deficits, and evaluating treatment outcomes as also being important in aiding with differential diagnosis (Matson, Nebel-Schwalm, & Matson, 2007).

One particular component in the assessment of an ASD that has been shown to be particularly effective in differential diagnosis is obtaining a child's developmental history (Deprey, & Ozonoff, 2009). In reviewing the child's developmental history, looking into the stability of behaviors over time and the presence of the symptoms across different contexts proves beneficial. The more consistent the behaviors (e.g., difficulties in social interaction,

poor eye contact) are across time and setting (e.g., school, home), the more likely the child is to have an ASD. For example, if a child demonstrates difficulties with social communication and repetitive behaviors consistently throughout his or her lifetime in multiple contexts, it would be more likely that the child has an ASD, as opposed to a disability that may have symptom overlap with an ASD (e.g., Intellectual Disability, Communication Disorder). Looking into the onset of the behaviors, changes or additions to baseline behaviors, and response to treatment also provides clues to aid in differential diagnosis (Deprey & Ozonoff, 2009; Lainhart, 1999).

Utility Indices of Behavior Checklists

An important quality for parent behavioral checklists to have that may identify children with possible ASDs, particularly those that may be used to aid in determining special education placement decisions, is sound utility. The utility of a measure is evaluated through examining a number of aspects such as, but not limited to, sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio.

Sensitivity and Specificity

Sensitivity and specificity are important to an instrument's utility. Sensitivity refers to the proportion of an instrument's ability to detect true positives in the population relative to the false positives. That is, the child is identified as having an ASD by the measure (i.e., ASRS), when the child does in fact meet criteria for the disability. Specificity refers to an instrument's ability to detect the proportion of true negatives relative to false negatives in a

population, or those children who do not meet criteria for ASDs on the measure and also do not have an ASD. Sensitivity and specificity are inversely related. As sensitivity is reduced, the instrument will under-identify the number of children with ASDs, whereas if specificity is reduced, the number of children with ASDs will be over-identified. Sensitivity and specificity are not influenced by sample size or base rates, but they are subject to sampling error and population differences Sensitivity and specificity correspond with Type I and Type II error. Type I error refers to the rate of false positives. Type II error refers to the rate of false negatives. The concepts of false positives and false negatives as well as sensitivity and specificity and their relation are illustrated in Figure 2.1.

Predictive Value

Predictive value also contributes to a diagnostic instrument's utility. Predictive value, in this case, would refer to the probability of the child having an ASD or not given the results of the ASRS. Positive predictive value is the probability of the presence of an ASD given that the ASRS indicates the child has the disorder, whereas negative predictive value indicates the probability of no presence of an ASD, given that the ASRS indicates the child does not have the disorder. Predictive value is influenced by the sensitivity and specificity of an instrument as well as the prevalence of a condition, in this case ASDs. As sensitivity increases, a negative result means it is less probable the individual has the disorder, indicating greater negative predictive value (i.e., negative test result rules out disorder). As specificity increases, a positive result means it is more probable the individual has the disorder, indicating greater positive predictive value (i.e., positive test result confirms disorder is present) (Essex-Sorlie, 1995).

		Predictor (ASRS)	
		No ASD	ASD
Criterion (ADOS)	ASD	I False Negative	II True Positive
	No ASD	III True Negative	IV False Positive

Sensitivity =II/(II + I)
Specificity = III/(IV + III)

Figure 2.1 *Illustration of False Positive, False Negative, Sensitivity and Specificity*

In addition to sensitivity and specificity, base rates of a condition also influence predictive power. In contexts where base rates are at the extremes of low or high, greater rates of sensitivity and specificity may still lead to high rates of false positives and false negatives For example, with a highly specific test, in contexts where base rates are low, positive results have a high chance of being false positives. ASDs have a low base rate. Therefore, any instrument set out to detect ASDs will have a high number of false positives. In contexts with high base rates, a highly sensitive test with negative results will result in a large number of false negatives.

Likelihood Ratio

Likelihood ratios are another way to analyze the utility of an instrument. A likelihood ratio expresses the likelihood that a given result would be expected in an individual with the condition compared to the likelihood that the same result would be expected in the individual without the condition (Fletcher & Fletcher, 2005). As an example, if, when examining the ASRS, it was found that 90% of children who had an ASD had a score of 75 on the ASRS (meaning sensitivity at 90%) and that 10% of children had other conditions (e.g., speech and language delay), a score of 75 would be 9 times more likely (i.e., 90/10) in a child with an ASD.

Likelihood ratios come in positive or negative forms. Positive likelihood ratio (LR+) is the portion of individuals with a positive test result, or those that the ASRS indicates as having an ASD (sensitivity), to the proportion of those who have a positive result, but do not have an ASD (1 – specificity) (Fletcher & Fletcher, 2005). The negative likelihood ratio (LR-) is the proportion of individuals who have a negative result, or those that the ASRS indicates are not on the autism spectrum, but who in fact have an ASD (1 – sensitivity) to the proportion of those who have a negative result, or the ASRS indicates they do not have an ASD, and, in fact, they do not (specificity). For an instrument to be a significant predictor LR+ values should be greater than 1 and LR- values should be a positive fraction between 0 and 1. Likelihood ratios equal to 1 indicate the test results are no more likely in those with the condition than without and when values drop below 1 for LR+ and above 1 for LR-, test results would mean the opposite (Spitalnic, 2004). A likelihood ratio around 1 typically has little value. Likelihood ratios with values greater than 1 indicate that the disorder, in this case

an ASD, is likely to be present, with higher numbers indicating a higher likelihood that an individual would have an ASD (McGee, 2002).

Cut Scores

Cutoff scores are another component of a behavioral measure contributing to its utility. A cutoff score on the ASRS can be defined as the minimum score the measure utilizes to indicate a child may have an ASD. For the ASRS, Goldstein and Naglieri (2010), in their standardization sample, found an average T-score of 72 on the Total Score when identifying children who may have an ASD. This T-score can be thought of as an ideal cut score in the identification of children with potential ASDs.

Issues of sensitivity and specificity are generally considered to be of limited value and difficult to interpret without examining them alongside cutoff scores (Johnson, Jenkins, & Petscher, 2009). This is because where the cutoff score is established will affect the values of sensitivity and specificity. Researchers must decide which is more important with any given measure, over-identifying or under-identifying individuals (in this case children with ASDs), and determine the optimal cutoff score based on that decision.

Receiver Operating Characteristic Curve

The Receiver Operating Characteristic (ROC) Curve (Metz, 1978) provides a test of the utility of a measure by examining how sensitivity and specificity change with specific cutoff scores. Used in conjunction with logistic regression, the ROC curve allows an examination of the differences in cutoff scores and their subsequent effect on the sensitivity, specificity and predictive validity of a particular measure. The ROC curve is represented in a

graphic format with specificity graphed on the X-axis and 1- sensitivity graphed on the Y-axis. In this manner, the graph can represent the optimal cut score and corresponding best possible sensitivity or specificity for the measure.

The primary meaning in the ROC curve analysis derives from the area under the curve (AUC). The AUC represents the ability of the instrument to distinguish between children with a particular classification and those without. The AUC coefficient can be interpreted as the percent of cases where a higher probability is assigned to a correct case (i.e., child with an ASD) as opposed to an incorrect case (i.e., child without an ASD), or how well a parameter (cutoff score) can distinguish between the two groups. A coefficient of .5 describing the AUC indicates the measure is no better than chance. The closer the coefficient is to +1, the better the measure is able to distinguish between two groups. As an example, an AUC of .95 would indicate the ASRS has a 95% chance of distinguishing between the child who has an ASD and another condition. In the ROC Curve, the point of the curve that comes closest to the upper left hand corner is considered the point at which sensitivity and specificity are maximized and a corresponding optimal cut score can be determined.

When examining the aforementioned psychometric properties of an instrument, it should be noted that base rates significantly affect the utility analysis of any psychological measure. When base rates deviate from .50 into extreme highs and lows, statistical procedures become compromised (Douglas, Otto, & Borum, 2003). Cut scores for psychological measures are interrelated with base rates. Most clinical disorders have a less than 5% prevalence rate, meaning that if no measure is given, 95% of cases could be classified correctly if the result was “negative” (McFall, 2005). To this end, by taking the

base rates of ASDs into account across specific contexts (e.g., special education referrals), ASRS scores will result in more meaningful utility analysis.

Statement of Problem

The Autism Spectrum Rating Scales (ASRS; Goldstein & Naglieri, 2010) is a newly developed behavior measure that attempts to identify children with ASDs and distinguish this group from other populations of children with disabilities. The ASRS is intended for use in various grade levels, including preschool and K-12. However, the psychometric properties of the measure have not been studied independently of the work done during the developmental stages of the instrument. Certain areas of the ASRS need further exploration. However, to date the only work published is that conducted with the standardization sample.

Use with young children in special education settings. The examination of the ASRS with the standardization sample as presented by Goldstein and Nagliari (2010) does not indicate how the measure may distinguish between groups of children with developmental disabilities within a preschool population referred for possible special education services. This is an important and unique setting that will need further examination, as there is a need in the schools to be able to identify children early on who may have an ASD from children with other developmental disorders having a similar behavioral presentation. As resources in the schools are becoming increasingly scarce, a measure is needed in the school context that a professional may use without the need for extensive time and training, one that is able to differentiate between children with an ASD and children with other forms of developmental disabilities. Finally, the behavioral presentation of children in the preschool setting can often have overlap across disabilities, thus, a measure that is able to distinguish between the

different types of disorders typically found in that setting will be crucial to helping with early identification and determining the most appropriate interventions.

Unclear clinical sample from development literature. Another limitation of the standardization sample as reported in Goldstein and Naglieri (2010) is that almost half of the clinical sample used to differentiate children with ASDs is not clearly defined by the authors of the instrument. Similar to that which would be found in a special education setting, the ASRS standardization sample includes children with communication and cognitive delays, but the rest of the comparison sample is not clearly defined (e.g., only defined as "Other"), so it is not certain if the remaining children in the comparison sample are those who would be reflective of the disorders and delays found in the preschool special education context (e.g., general developmental delays) or whether they are more similar to groups of children that would be found in other settings (e.g., psychiatric diagnoses). Table 3.1 outlines the exact number of children in the ASRS clinical standardization sample, including those with an ASD and those with another diagnosis.

Psychometric properties. Another strong component of the ASRS that warrants further exploration of the measure with a special education population are the purported psychometric properties of the instrument as described to date by Goldstein and Naglieri (2010). There are several psychometric properties that contribute to the overall clinical utility of a measure such as the ASRS, as well as other published measures that attempt to identify ASDs and differentiate them from other types of disorders.

The ASRS standardization sample reports sound utility across a number of psychometric indices (e.g., sensitivity, specificity), outlined in Table 1.1. Of the current

Table 3.1: *Diagnostic Categories of the Clinical Standardization Sample of the ASRS*

Diagnosis	Parent	Teacher
Autism Spectrum Disorder	135	124
Communication Disorders	35	38
Delayed Cognitive Development	41	43
Other	58	69
Total	269	274

Note: ASRS = Autism Spectrum Rating Scales; Goldstein & Naglieri, 2010

measures used to identify children who may have an ASD, the ASRS appears to have comparable or better psychometric properties than those in use. More specific psychometric properties are outlined in detail in the Method section.

Research Questions

Six primary measures of classification accuracy will be used in the evaluation of the Autism Spectrum Rating Scale: sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio

Primary Research Questions

Research Question 1: What do sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio indicate about the validity of the ASRS when used with the recommended cut score in a preschool special education context?

Research Question 2: How do sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio change with different cut scores on the ASRS, and does this support using a different cut score with a preschool special education population?

Research Question 3: What is the AUC at the cutoff score of the ASRS that will optimize sensitivity and specificity for a preschool special education population?

Research Question 4: How would positive predictive value and negative predictive value (and possible cut score preferences) be expected to change for the ASRS with different base rates that might be found in different settings and applications?

Research Question 5: How would preferred cut scores be expected to change for applications of the ASR that differ in the perceived costs of false positive and false negative misclassifications?

Supplemental Research Questions

Supplemental Research Question 1: Are there significant differences in terms of ASRS cut scores depending on the participant's IQ, language ability, and adaptive skills?

Supplemental Research Question 2: Which specific items from the ASRS discriminate most between children with an ASD and children with other disabilities?

CHAPTER II

METHOD

Participants

Participants for the current study were recruited from a population of preschool children, between the ages of 2 years and 6 years of age, referred for assessment because of a possible developmental disability. Families were recruited from a public school district in the Western United States. The participants included in the current study are reasonably representative of the typical population of children referred for the identification and determination of eligibility for special education services because of a possible developmental disability. All aspects of the current study were approved by the University of Utah Institutional Review Board on March 9, 2011 (IRB # 00045549).

Participant Recruitment Procedures

The following steps were used to recruit participants for the current study:

Step 1: The primary investigator was given the names of all preschool children who were referred for assessment for special education eligibility determination. These children either failed a developmental screening measure (i.e., Developmental Indicators for the Assessment of Learning – Third Edition) given to all preschoolers enrolled in the district, had been identified through the State of Utah's Child Find program (i.e., IDEA 2004, Section B), or were referred by a school's special education team as having a suspected developmental

disability. Among other types of developmental disabilities, referred children may have had a Speech or Language Impairment, Intellectual Disability, and/or Autism (IDEA, 2004). (See Appendix B for a complete summary of IDEA disability criteria.)

Step 2: The primary investigator then contacted parents of the preschool children by email to determine if they were interested in participating in the current study (see Appendix D for information that parents were provided). After the initial e-mail was sent, a follow-up telephone call was then placed to confirm interest of the parents in participating in the study.

Step 3: Upon indicating interest in participating, parents were asked to schedule a time to meet with the primary investigator to review procedures being used in the study, the safeguards used to ensure confidentiality and to protect their child's identity in the study, and sign a written consent form (see Appendix E for consent form). Upon receipt of the signed consent form, the investigator scheduled a time to meet and conduct the assessments with the caregiver and child. Sixty-eight families were identified as possible participants in the study. One family refused because they were unable to attend the appointment required for the study.

Participants for the current study (N = 67) were divided into two groups, children with an autism spectrum disorder (ASD) (n = 37) and those with a general developmental disability (DD) (n = 30). Table 4.2 outlines key demographic variables between the two groups. The ASD group consisted of 37 children, mean age 3.30 years, SD = 0.85, of which 32 were males and 5 were female. Within the ASD group, White participants comprised 89% of the sample, with non-Whites making up 11% of the group. The DD group consisted of 30 children, mean age 4.22 years, SD = .93, of which 23 were males and 7 were female. Regarding ethnicity in the DD group, 87% were White and 13% were non-White.

Table 4.2: *Sample Characteristics*

Demographic	ASD	DD
	(*n* = 37)	(*n* = 30)
Age; Mean (Standard Deviation)	3.30 (0.85)	4.22 (0.93)
Sex		
Male	32	23
Female	5	7
Ethnicity		
American Indian/Alaska Native	2	0
Asian	0	0
Black or African American	0	0
Hispanic/Latino	1	4
Native Hawaiian/Pacific Islander	1	0
White	33	26

Participants were recruited from a pool of children who were referred for a special education eligibility evaluation so, they subsequently received a special education classification utilizing current IDEA rules and regulations by the school district after their assessment. Participants in the current study met criteria for the following special education classifications: Autism (*n* = 28), Speech and Language Impaired (*n* = 5), Developmental Delay (*n* = 28), Multiple Disabilities (*n* = 1), and Other Health Impaired (*n* = 1). The

remaining participants (n = 4) were found not eligible for special education services after the special education evaluation.

Exclusionary Criteria

There were no exclusionary criteria established for any of the families and children. As long as the children were identified by the previously mentioned means (e.g., failing DIAL-3 screening), they were deemed eligible for the current study. Of the total families that were contacted to participate, one family declined (1.5% of the total recruited), indicating that they were too busy to participate.

Instrumentation and Measures

Autism Spectrum Disorder

Autism Spectrum Rating Scales. The Autism Spectrum Rating Scales (ASRS) is a norm-referenced behavioral measure designed to identify symptoms, behaviors, and associated features of ASDs. The ASRS comes in two forms, one for children 2 to 5 years of age and the other for ages 6 to 18. The current project utilized the form for the 2- to 5-year-olds. Scoring for the ASRS yields a Total Score, which is considered to be the most broadband measure of ASD symptoms. It is comprised of two subscales, Social/Communication and Unusual Behaviors. All scales of the ASRS are T-scores. Three sample groups were used in the standardization of the ASRS: children with an ASD, another clinical sample, and a sample from the general population. The clinic sample was comprised of children with diagnoses such as Delayed Cognitive Development and Delayed

Communication Development. T-scores of the groups' means are presented in Table 5.2. There was a significant difference in scores between the ASD versus General Population groups, as well as the ASD versus the Other Clinical Sample groups, on all three scales for the standardization sample (i.e., Total Score, Social/Communication, and Unusual Behaviors).

Currently, there are no independent empirical studies examining the reliability and validity of the ASRS. Studies performed on the standardization sample indicate that the ASRS has sound reliability. Internal consistency of the ASRS Total Score was .97 and coefficients for the Index scores ranged from .85 (Unusual Behaviors Index) to .98 (Social/Communication). Table 6.2 outlines the reliability of the ASRS with the standardization sample.

The standardization research examined the relationship between the ASRS and three other commonly used measures of ASDs: the GARS-2, the GADS, and the CARS. The ASRS was moderately correlated with the GARS (r = .80) and the GADS (r = .71), but correlation was lower for the CARS (r = .36). Table 7.2 outlines validity data for the standardization sample of the ASRS.

Table 5.2: *Parent ASRS T-score Comparison by Sample and Scale as Reported in the Standardization Sample*

	Total Score	Social / Communication	Unusual Behaviors
ASD Sample	72.8	70.4	68.6
General Sample	47.7	46.6	48.8
Other Clinical Sample	46.9	49.8	43.7

Note: ASRS = Autism Spectrum Rating Scales

Table 6.2: *Reliability of the ASRS Standardization Sample*

	Parent	Teacher
Scale	Cronbach's Alpha	Cronbach's Alpha
ASRS Total Score	.97	.97
Social/Communication	.96	.97
Unusual Behavior	.94	.93
DSM-IV-TR Scale	.95	.95
Peer Socialization	.89	.91
Adult Socialization	.77	.82
Social/Emotional Reciprocity	.91	.93
Atypical Language	.74	.70
Stereotypy	.81	.78
Behavioral Rigidity	.90	.90
Sensory Sensitivity	.81	.78
Attention/Self-Regulation	.86	.86

Note: ASRS = Autism Spectrum Rating Scales

Additional psychometric properties of the standardization sample were excellent, demonstrating that the ASRS is a sound measure for assessing ASDs. Sensitivity and specificity statistics were examined as part of the standardization sample.

Sensitivity for the Total and Index Scores ranged from 95.0% to 89.8%, while specificity for the Total and Index Scores ranged from 94.7% to 90.3%. Positive predictive power was 91.3% and negative predictive power, 88.7%. False positives (or Type I error)

Table 7.2: *Validity of the ASRS (2-5) Standardization Sample*

	Parent ASRS		Teacher ASRS	
	Obtained *r*	Corrected *r*	Obtained *r*	Corrected *r*
GARS	.83	.61	.76	.41
GADS	.63	.49	.76	.56
CARS	.06	.06	.50	.66

Note: Validity is reported on the ASRS Form for ages 2 – 5 years. ASRS = Autism Spectrum Rating Scales; GARS = GARS Autism Index; GADS = GADS Asperger's Disorder Quotient; CARS = CARS Total Raw Score. All correlations significant, $p < .01$, except for the ASRS Parent correlation with the CARS. Corrected *r* corrected for measurement error.

were 9.7%, whereas false negatives (or Type II error) were 10.2%. Percentages are based on a comparison of children with ASD in the standardization sample to a neurotypical population.

Autism Diagnostic Observation Schedule (Lord, Rutter, DiLavore, & Risi, 1999). The Autism Diagnostic Observation Schedule (ADOS) is a semi-structured behavioral measure designed to assess children or adults suspected of having an ASD. Activities and interviews of the ADOS are designed to elicit responses in the area of social and reciprocal communication that may be typical of children with ASDs. There are four different modules of the ADOS that may be given, with each module based on the child's language ability and age. The ADOS typically takes 45-60 minutes to complete. The ADOS has been demonstrated to have sound psychometric properties (Lord et al., 1999).

Language

Peabody Picture Vocabulary Test – Fourth Edition (Dunn & Dunn, 2006). The Peabody Picture Vocabulary Test – Fourth Edition (PPVT-4) is a recently published standardized measure of receptive vocabulary. During the assessment, individuals are asked

to identify a picture that represents a word stated by the administrator by selecting one picture out of a series of four. Data from the normative sample, which included children with developmental disabilities, have demonstrated that the PPVT-4 has sound reliability and validity (Dunn & Dunn, 2006). Reliability coefficients have been found to be between .89 and .95 when examining internal consistency, alternate form and test-retest reliability. Validity studies have shown the PPVT-4 to correlate with other known well-established expressive language measures (e.g., Clinical Evaluation of Language Functions – Fourth Edition; CELF-4) with coefficients ranging from .72 to .84.

Expressive Vocabulary Test – Second Edition (Williams, 2007). The Expressive Vocabulary Test – Second Edition (EVT-2) is a standardized measure of expressive vocabulary and word retrieval. Individuals are presented with a stimulus picture and asked to answer a question by the examiner, verbally give the most appropriate label, or provide a synonym for the stimulus. Studies demonstrate that the EVT-2 has sound reliability and validity (Williams, 2007). The normative sample included a population matched to the 2004 Census data, including representative samples of children with developmental disabilities. Reliability coefficients for internal consistency, alternate form, and test-retest reliability range from .87 to .93. Validity coefficients for the EVT-2 range from .77 to .81 when compared with concurrent measures (i.e., CELF-4).

Intelligence

Stanford-Binet Intelligence Scales – Fifth Edition (Roid, 2003a). The Stanford-Binet Intelligence Scales – Fifth Edition (SB5) is an individually administered intelligence test for individuals 2 – 85 years of age. The SB5 assesses five domains: Fluid Reasoning (Verbal

and Non-Verbal problem solving using reasoning), Knowledge (fund of general information), Quantitative Reasoning (working and solving problems with numbers), Visual-Spatial Processing (analyzing patterns, relationships, and spatial orientation), and Working Memory (ability to store, sort and transform information in short-term memory). These domains are aggregated into Verbal and Non-Verbal and Full Scale IQ components. Scoring is based on standard scores with a mean of 100 and a standard deviation of 15. The SB5 has sound psychometric properties. Reliability coefficients range from .97 to .98 for the Full Scale IQ score (Roid, 2003b). Validity studies (e.g., construct, criterion, concurrent) also indicate that the SB5 is a psychometrically sound instrument (Roid, 2003b). For the purposes of the study, the Abbreviated version of the SB5 was utilized.

Leiter International Performance Scale – Revised (Roid & Miller, 1997). The Leiter International Performance Scale – Revised (Leiter-R) is a nonverbal individually administered assessment of nonverbal intelligence designed to be used with individuals 2 – 20 years of age. The Leiter-R consists of subtests scores that yield scaled scores based on a mean of 10 and a standard deviation of 3, and an intelligence quotient based on a mean of 100 and a standard deviation of 15. During the administration, the examinee is asked to solve a series of tasks that do not require verbal responses. Subtests of the Leiter-R consist of Figure Ground, Form Completion, Repeated Patterns and Sequential Order. Instead, examinee's utilize a series of manipulatives and cards. The Leiter-R has demonstrated sound reliability and validity (Roid & Miller, 1997). For the current study, the Brief IQ was utilized.

Adaptive Measure

Scales of Independent Behavior – Revised (Bruininks, Woodcock, Weatherman, & Hill, 1996). The Scales of Independent Behavior – Revised (SIB-R) is a norm-referenced measure that assesses an individual's adaptive behavior across several domains within the contexts of the home and community. The SIB-R yields adaptive scores in the following areas: Broad Independence, Social Interaction and Communication Skills, Personal Living Skills, Community Living Skills, Motor Skills, and a General Maladaptive Index. The SIB-R also examines problematic behavior and provides scores in the following areas: Internalized Maladaptive Index, Asocial Maladaptive Index, and Externalized Maladaptive Index. The SIB-R is administered to parents in a checklist format. The SIB-R has shown to have sound reliability and validity (Strauss, Sherman, & Spreen, 2006).

Data Collection Procedures

Data collection took place over the course of 1 year, from March 2011 to March 2012. The data for the current study's assessment battery were collected several sessions per participant lasting a cumulative total of 3-4 hours. Most of this time was spent in face-to-face sessions with each child while parents completed checklists. Each group of sessions occurred within a 45-day timeframe, which corresponds to the deadline within which an IDEA evaluation must be completed. Detailed steps of the data collection procedure are outlined below.

Step 1: Each child was scheduled to participate in testing with the ADOS by the primary investigator, a licensed school psychologist, who has completed formal ADOS

training through Western Psychological Services (WPS), together with another similarly-trained school psychologist intern, or licensed professional.

Step 2: At the time of the ADOS testing, parents were asked to complete the Parent Ratings version of the ASRS and the Scales of Independent Behavior – Revised.

Step 3: If the children were currently enrolled in a preschool, the primary investigator contacted the child's preschool teacher, who was then asked to complete the Teacher/Childcare Provider Ratings version of the ASRS form.

Step 4: If the student was not currently enrolled in a preschool, the parents were asked to have a daycare provider complete the Teacher/Childcare Provider Ratings version of the ASRS.

Step 5: After the ADOS administration, all children completed the Peabody Picture Vocabulary Test – Fourth Edition and the Expressive Vocabulary Test – Second Edition. Based on the results of the language testing, a verbal or nonverbal IQ test was administered. A verbal IQ test was administered to children whose language scores fell within the average range. A nonverbal IQ test was administered to children whose language scores fell within the low average range or lower. This algorithm was used so that potential language difficulties did not interfere with establishing a valid IQ score for a child.

Step 6: Children who had language scores in the average range were administered the Stanford-Binet Intelligence Scales – Fifth Edition. Those children whose language scores fell within the low average range or lower were administered the Leiter International Performance Scale – Revised.

Step 7: Based on test results, children were either identified as having an ASD or a general DD. Decisions regarding group membership (ASD vs. DD) were based on the following:

i. The ADOS was used as the criterion to classify children into either the ASD or DD groups. The ADOS Total Score is comprised of the Communication + Reciprocal Interaction subtests. On the ADOS, a Total Score of $\geq$ 7 on Modules 1 and 3, or $\geq$ 8 on Module 2 signifies the presence of an ASD. Children were included in the ASD group if they met criteria on the ADOS, regardless of whether they had another disability (e.g., Intellectual Disability). No other measures besides the ADOS were used to classify the children into the ASD or DD groups.

ii. The DD group consisted of children who fell below the cut score for an ASD on the ADOS Total score. Children in the DD group had delays in one or more areas of development and often met the IDEA special education eligibility criteria for classifications other than Autism, such as Intellectual Disability (ID) and Speech and Language Impairment (SLI) (see Appendix C for complete summary of IDEA Classification Criteria).

Analysis

Sensitivity for the ASRS was calculated by dividing the true positives (TP) by the true positives plus false negatives (FN); that is, sensitivity = TP / (TP + FN). True Positives (TP) were those participants whose scores fell within the clinically significant range on the ASRS Total Score (T-score $\geq$ 70) and who were determined to be in the ASD group using the

ADOS Total Score. The False Negatives were those participants whose scores did not fall within the clinically significant range on the Total Score scale of the ASRS (T-score $\leq$ 69), but who were identified as being in the ASD group using the ADOS Total Score.

Specificity was calculated by dividing the true negatives (TN) by the false positives (FP) plus true negatives (TN); that is, specificity = TN/ (TN + FP). True Negatives were those participants who were not in the significant range on the Total Score of the ASRS, and were determined to be in the DD group using the ADOS Total Score. False Positives were those participants who were in the significant range on the Total Score of the ASRS but were not considered to have an ASD based on the ADOS Total Score.

The positive predictive value of the ASRS was determined by the following formula: True Positives / (True Positives + False Positives). The negative predictive power of the ASRS was determined by the following formula: True Negatives / (True Negatives + False Negatives). The positive likelihood ratio was determined by the following formula, [Sensitivity / (1 - Specificity)]. The negative likelihood ratio was determined by the following formula: [(1 – Sensitivity) / Specificity]. A Receiver Operating Characteristic (ROC) curve was utilized to examine the AUC at the optimal cutoff score that resulted in the most psychometrically sound range of scores for sensitivity and specificity on the ASRS by mapping sensitivity on the y-axis and 1 – specificity on the x-axis.

Scatter plots were utilized to map out IQ, language, and adaptive scores in an exploratory manner and error rates were calculated by developmental area. A point-biserial correlation was used to determine what items discriminated most between children on the autism spectrum and children with other disabilities (ASD vs. DD).

CHAPTER III

RESULTS

General Descriptive Statistics

When examining demographic variables, the ASD and DD groups did not differ with regards to sex, χ^2 (1, N = 67) = 1.09, $p > .05$, and ethnicity, χ^2 (3, N = 67) = 4.95, $p > .05$, but a significant difference was found with age, t (65) = 4.22, $p < .001$. As a result of the significant age difference between the ASD and DD groups, subsequent analysis where age may be a factor affecting outcome, appropriate measures were taken to compensate. For example, age was introduced as a covariate in the analysis when necessary.

The ADOS was used as the primary criterion measure for the current study. Table 8.3 outlines group means for the ASD and DD groups as reflected in the ADOS scales. On the ADOS Total scale, the ASD group had a mean of 12.27, SD = 3.84, and the DD group had a mean of 2.77, SD = 2.23. The ASD group means for the ADOS Total scale are within the range expected for someone with an ASD, whereas those in the DD group fall well below the threshold for an ASD.

Correlations

The correlation between the Parent and Teacher ASRS Total scale scores was computed and found to be not statistically significant. Subsequent correlation analysis was

Table 8.3: *Means and Standard Deviations of ADOS Scales by Group*

Scale	ASD	DD
	(*n* = 37)	(*n* = 30)
ADOS Total*	12.27 (3.84)	2.77 (2.23)
Communication	4.57 (1.28)	3.17 (11.54)
Reciprocal Social Interaction	7.70 (3.11)	1.73 (1.87)
Play	1.70 (1.50)	0.57 (0.67)
Stereotyped Behaviors/ Repetitive Interests	1.70 (1.86)	0.33 (0.76)

Note: ADOS = Autism Diagnostic Observation Schedule; ASD = Autism Spectrum Disorder; DD = Developmental Delay
*Scores indicating presence of an ASD are $\geq$ 7 for ADOS Modules 1 and 3, $\geq$ 8 for ADOS Module 2

performed to examine the ASRS and ADOS scales together. Because the Parent and Teacher ASRS Total scale scores were not correlated, each measure was used in a separate correlation analysis with the ADOS scales. Tables 9.3 and 10.3 outline the correlations between the Total aggregated scales and the subscales for the ASRS and ADOS measures. Across measures, the Parent ASRS Total Score was not correlated with any of the ADOS scales, including the ADOS Total scale, the ADOS Communication subscale, or the ADOS Reciprocal Social Interaction subscale. When comparing ASRS with ADOS scales, significant correlations were found with the Parent ASRS Social/Communication subscale and the ADOS Total scale and the ADOS Reciprocal Social Interaction scale. The Parent ASRS Unusual Behaviors subscale was not correlated with any ADOS scales.

Table 9.3: *Parent ASRS and ADOS Scale Correlations*

Subtest	1	2	3	4	5	6
1. ASRS Total	-	.83***	0.86***	0.23	0.06	0.23
2. ASRS Social/ Communication		-	0.51***	0.46***	0.05	0.44***
3. ASRS Unusual Behaviors			-	0.02	0.05	0.05
4. ADOS Total				-	0.14	0.97***
5. ADOS Communication					-	0.14
6. ADOS RSI						-

Note: ASRS = Autism Spectrum Rating Scales; ADOS = Autism Diagnostic Observation Schedule; RSI = Reciprocal Social Interaction.
*$p < .05$
**$p < .01$
***$p < .001$

When examining results for the Teacher ASRS, significant correlations were found between the Teacher ASRS Total scale and the ADOS Total scale as well as the ADOS Reciprocal Social Interaction scale. When comparing the Teacher ASRS and ADOS scales, other significant correlations were found between the Teacher ASRS Social/Communication scale, the ADOS Total scale, and the ADOS Reciprocal Social Interaction scale. No significant correlations were found between the Teacher ASRS Unusual Behavior scale and any of the ADOS scales. Across Parent and Teacher ASRS, the highest frequency and greatest magnitude of correlations were found between the ADOS scales and the ASRS Social/Communication subscale.

Table 10.3: *Teacher ASRS and ADOS Scale Correlations*

Subtest	1	2	3	4	5	6
1. ASRS Total	-	.73***	0.82***	0.29*	0.09	0.32*
2. ASRS Social/ Communication		-	0.38**	0.39***	0.04	0.39**
3. ASRS Unusual Behaviors			-	0.06	0.06	0.14
4. ADOS Total				-	0.14	0.97***
5. ADOS Communication					-	0.14
6. ADOS RSI						-

Note: ASRS = Autism Spectrum Rating Scales; ADOS = Autism Diagnostic Observation Schedule; RSI = Reciprocal Social Interaction.
*$p < .05$
**$p < .01$
***$p < .001$

Multivariate Analysis of Variance

The ASRS was used in the current study as the experimental measure, or predictor. A MANOVA was used to compare means of the Parent and Teacher ASRS scales for the ASD. and DD groups to determine the instrument's ability to classify children with ASDs correctly Appendix F outlines the means and standard deviations of the Parent and Teacher ASRS. Only those subscales with the least amount of overlap were entered into the model. The scales with the least amount of overlap were the Treatment Scales, which consist of: Peer Socialization, Adult Socialization, Social/Emotional Reciprocity, Atypical Language,

Stereotypy, Behavioral Rigidity, Sensory Sensitivity, and Attention/Self-Regulation. The ASRS subscales (Unusual Behaviors, Communication, and DSM-IV-TR) and aggregated scale (ASRS Total Score) contain the items of, and thus, overlap with, the Treatment scales. For the Parent ASRS, the multivariate effect was significant for group, F (8, 58) = 2.27, $p <$ .05, indicating a significant ability for the Parent ASRS to classify children as ASD. The univariate F tests showed a significant difference between the ASD and DD groups for Peer Socialization, F (1, 65) = 8.77. For the Teacher ASRS, the multivariate effect was also significant for group, F (8, 44) = 2.20, $p < .05$. For the Teacher ASRS, the univariate F tests showed a significant difference between the ASD and DD groups for Peer Socialization, F (1, 51) = 4.01, $p < .05$, and Social/Emotional Reciprocity, F (1, 51) = 5.05, $p < .05$. Given there was a statistically significant difference in age between the ASD and DD groups, the same MANOVA was performed with age as a covariate. Entering age as a covariate in the MANOVA, the test was no longer significant for the Parent ASRS, F (8, 57) = 1.32, $p > .05$ or Teacher ASRS, F (8, 43) = 1.39, $p > .05$.

Figures 3.3 and 4.3 illustrate the means in T-scores of the ASD and DD groups for the Parent and Teacher ASRS versions, respectively. On each figure, the dark bar indicates the ASD group and the light bar reflects the DD group. Overall, means for the ASD group tended to be elevated on a majority of the scales of the ASRS when compared to the DD group.

T-Tests

In addition to the measures of autism, the current study also assessed several domains of developmental functioning, including intelligence, adaptive skills, and language. The Leiter-R or Stanford-Binet – Fifth Edition were administered as assessments of intellectual

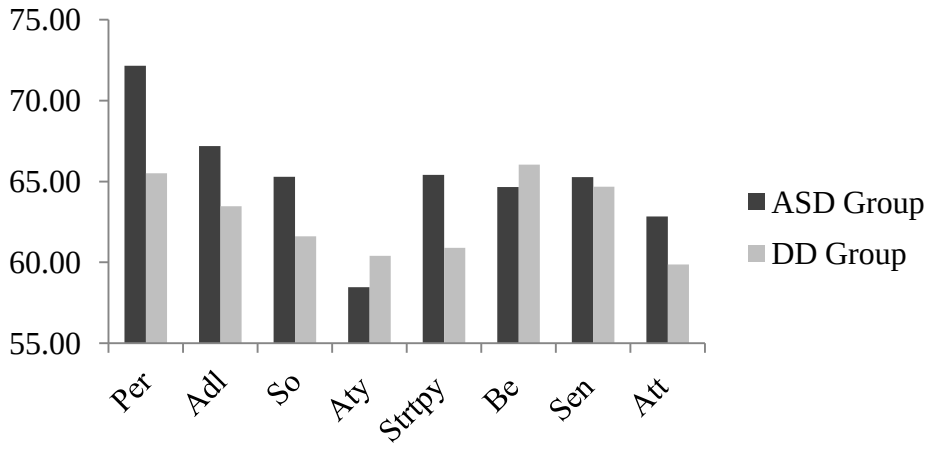

Figure 3.3 *T-scores of the ASRS, Parent Version Scales by Group*

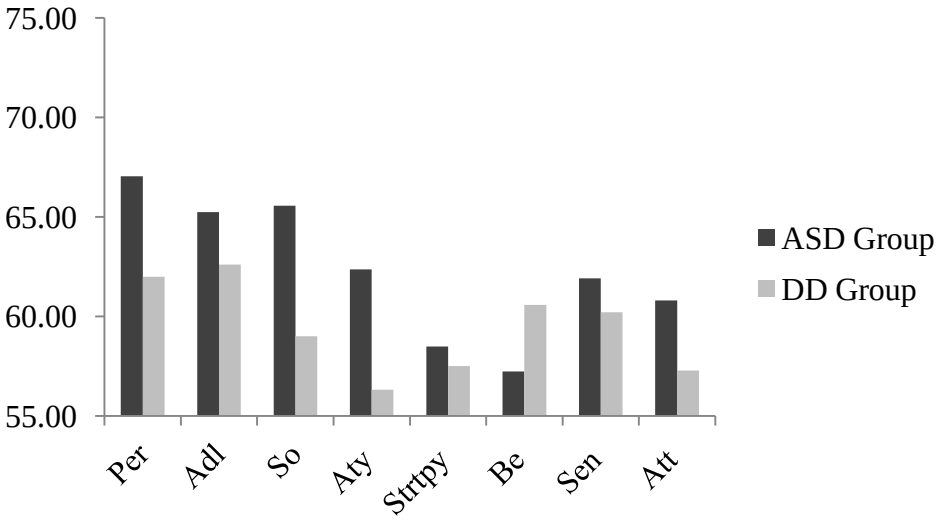

Figure 4.3 *T-scores of the ASRS, Teacher Version Scales by Group*

(Soc = Social/Communication, Un = Unusual Behaviors, DSM = DSM-IV-TR, Per = Peer Socialization, Adl = Adult Socialization, So = Social/Emotional Reciprocity, Aty = Atypical Language, Strtpy = Stereotypy , Be = Behavioral Rigidity, Sen = Sensory Sensitivity, Att = Attention/Self-Regulation) by group.

abilities, depending on the child's language scores. Children who tested within the average range for language were administered the Stanford-Binet – Fifth Edition, whereas those with language abilities in the low average range or below were administered the Leiter-R, which was nonverbal. Adaptive abilities were assessed using the Scales of Independent Behavior – Revised. Receptive vocabulary was measured with the Peabody Picture Vocabulary Test – Fourth Edition and expressive vocabulary was measured with the Expressive Vocabulary Test – Second Edition. Appendix G outlines the specific means and standard deviations for the developmental variables for the ASD and DD groups, reported as standard scores with a mean of 100 and a standard deviation of 15. When examining the developmental variables, significant differences between the ASD and DD groups were found in all areas: intellectual abilities $t(63) = -3.09$, $p < .01$, adaptive skills, $t\ (63) = -4.29$, $p < .001$, receptive vocabulary $t(59) = -4.02$, $p < .001$, and expressive vocabulary, $t(59) = -2.76$, $p > .01$

Figure 5.3 graphically depicts the contrast between the standard scores (mean = 100; standard deviation = 15) of the developmental variables for the ASD and DD groups. In the figure, the light bars represent the standard scores for the ASD group and the darker bar represents the DD group. Included in the graph are error bars that indicate a 95% confidence interval around the mean. For these particular variables, in a standardization sample, the mean is 100 and the standard deviation is 15. In general, the ASD groups tends to score lower on the variables than the DD group.

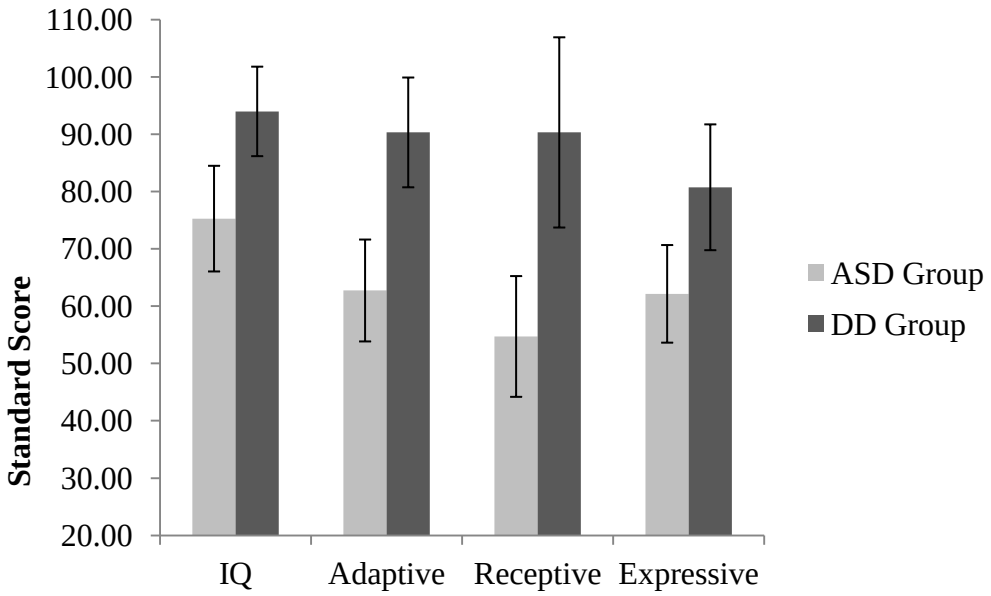

Figure 5.3 *Mean Standard Scores of the Developmental Variables by Group*

Error bars denote 95% confidence interval around the mean.

Research Question Analysis

Research Question 1

The first research question inquired into the meaning of the sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood variables for the validity of the ASRS when used with the recommended cut score in a population of preschool children referred for special education eligibility determination.

To explore this research question, the ASRS was utilized as the experimental measure in the identification of children on the autism spectrum, with the ADOS used at the criterion measure. Using a cut score that traditionally signifies clinical significance (T-score $\geq$ 70; 98th

percentile) as the recommended cut score, the aforementioned aspects of the clinical utility (e.g., sensitivity, specificity) of the ASRS were determined.

True positives for the Parent ASRS were 24 of 67 (36%), with true negatives at 19 of 67 (28%), creating a hit rate [(True Positive + True Negative)/All cases] for the ASRS of 64%. The Parent ASRS had a Type I (false positive) error rate of 11 of 67 (16%) and a Type II (false negative) error rate of 13 of 67 (19%). Figure 6.3 outlines the results of the Parent ASRS classification when compared to the current "gold standard," the ADOS, as the criterion in the classification of children with potential ASDs.

True positives for the Teacher ASRS were 20 of 53 (38%), with true negatives at 13 of 53 (25%), creating a hit rate [(True Positive + True Negative)/All cases] for the ASRS of 62%. The Teacher ASRS had a Type I (false positive) error rate of 8 of 53 (15%) and a Type II (false negative) error rate of 12 of 53 (23%). Figure 7.3 outlines the results of the Teacher ASRS classification when compared to the current "gold standard," the ADOS, as the criterion in the classification of children with potential ASDs. True positives for the Short Form ASRS were 19 of 58 (33%) with true negatives at 17 of 58 (29%), creating a hit rate [(True Positive + True Negative)/All cases] of the ASRS of 62%.The Short Form ASRS had a Type I (false positive) error rate of 10 of 58 (15%) and a Type II (false negative) error rate of 12 of 58 (21%)

When using the T-score of $\geq$ 70 as the cut score for the Parent ASRS, the measure's sensitivity was found to be 64.86, 95% CI [47.5, 78.9] and specificity was found to be 63.30, 95% CI [45.5, 79.0].

		Predictor[#]	
		No ASD	ASD
Criterion[*]	ASD	13	24
	No ASD	19	11

Figure 6.3 *Parent ASRS Hit Rate and Error Analysis with an ASRS Cut T-score of ≥ 70*

		Predictor[#]	
		No ASD	ASD
Criterion[*]	ASD	12	20
	No ASD	13	8

Figure 7.3 *Teacher ASRS Hit Rate and Error Analysis with an ASRS Cut T-score of ≥ 70*

(ASD = Autism Spectrum Disorder; DD = Developmental Delay; ASRS = Autism Spectrum Rating Scales; ADOS = Autism Diagnostic Observation Schedule). [#]Autism Spectrum Rating Scales, [*]Autism Diagnostic Observation Schedule

The positive predictive value for the ASRS was 68.6 and the negative predictive value was 59.4. The positive likelihood ratio was 1.77, with the negative likelihood ratio being .55. Using the cut score of ≥ 70 on the Teacher ASRS, sensitivity was found to be 52.00, 95% CI [31.3, 72.2], specificity 71.43, 95% CI [51.3, 86.8], positive predictive power 61.9, negative predictive power 62.5, positive likelihood ratio of 1.82, and a negative likelihood ratio of .67. Figure 8.3 illustrates, in a scatter plot form, the classification rate of the Parent and Teacher ASRS, as well as the false negative and false positive rates. Dotted lines indicate cut scores for ASD on each respective axis, with higher scores on each measure indicating greater likelihood of the child showing behaviors consistent with being on the autism spectrum.

Research Question 2

The aim of the second research question was to inquire into how sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio change with different cut scores. An additional purpose was to determine if the data may support using a different cut score with a population of preschool children referred for special education eligibility. Sensitivity, specificity, positive predictive power, negative predictive power, positive likelihood ratio, and negative likelihood ratio and each index's rate of change associated with different cut scores is examined in Table 11.3 for the Parent ASRS and Table 12.3 for the Teacher ASRS.

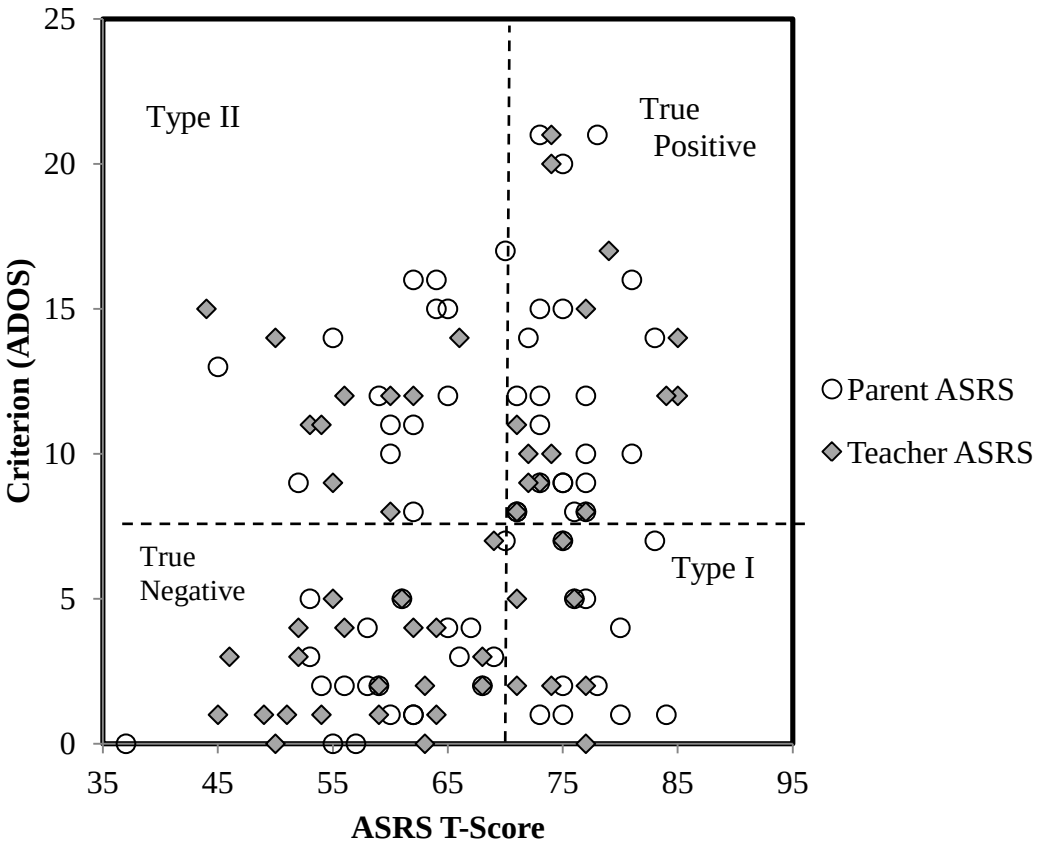

Figure 8.3 *Scatter Plot Illustrating Classification Rate and Error Rates of the Parent and Teacher ASRS Using a Cut Score of T ≥ 70*

Given the small sample size, cut scores could not be presented in a sequential format as the classification matrix did not always change when moving from one cut score to the next higher. Thus, skips in cut scores on the table represent ties. Moreover, the number do not follow typical trends (e.g., PPV continues to increase on the Parent ASRS) because of the same reason, small sample size and number combinations in the classification matrix. The Youden index (Youden, 1950) captures the optimal performance of a diagnostic test, such as the ASRS. The Youden index is defined as [(sensitivity) – (specificity)] -1. It provides a

Table 11.3: *Cut Score and Utility Analysis of the Parent ASRS*

Cut Score	Sensitivity	95% CI	Specificity	95% CI	PPV	NPV	LR+	LR-
60	83.78	68.0 - 93.8	36.67	19.9 - 56.1	62	64.7	1.32	0.44
61	83.78	68.0 - 93.8	40.00	22.7 - 59.4	63.3	66.7	1.4	0.41
62	75.68	58.8 - 88.2	46.67	28.3 - 65.7	63.6	60.9	1.42	0.52
64	70.27	53.0 - 84.1	46.67	28.3 - 65.7	61.9	56	1.32	0.64
65	64.86	47.5 - 79.8	50.00	31.3 - 68.7	61.5	53.6	1.3	0.7
69*	64.86	47.5 - 78.9	63.30	45.5 - 79.0	68.6	59.4	1.77	0.55
70	64.86	47.5 - 78.9	63.30	45.5 - 79.0	68.6	59.4	1.77	0.55
72	48.65	31.9 - 65.6	63.33	43.9 - 80.1	62.1	50	1.33	0.81
73	35.14	20.2 - 52.5	66.67	47.2 - 82.7	56.5	45.5	1.05	0.97
75	21.62	9.8 - 38.2	73.33	54.1 - 87.7	50	43.1	0.81	1.07
76	18.92	8.0 - 35.2	76.67	57.7 - 90.1	50	43.4	0.81	1.06

Note: Absent cut scores reflect ties. CI = Confidence Interval; PPV = Positive Predictive Value; NPV = Negative Predictive Value; LR+ = Positive Likelihood Ratio; LR- = Negative Likelihood Ratio.
* Highest Youden index

summary of the test's accuracy. Using the Youden index, a cut score of 69 appears to be the most optimal point to maximize sensitivity and specificity for the Parent ASRS, whereas for the Teacher ASRS, a cut score of 68 appears to be the optimal score for this population.

Positive likelihood ratios of greater than 1 indicate that a particular disorder or disease is likely to be present and negative likelihood ratios between 0 and 1 indicate the absence of a particular disorder or disease (McGee, 2002). In the case of the current study, the concern is the presence or absence of an ASD. Looking at the results of the utility

Table 12.3: C*ut Score and Utility Analysis of the Teacher ASRS*

Cut Score	Sensitivity	95% CI	Specificity	95% CI	PPV	NPV	LR+	LR-
60	64.00	42.5 - 82.0	42.86	24.5 - 62.8	50.0	57.1	1.12	.84
61	64.00	42.5 – 82.0	46.43	27.5 - 66.1	51.6	59.1	1.19	0.78
62	60.00	38.7-78.9	50.00	30.6-69.4	51.7	58.3	1.20	.80
64	60.00	38.7-78.9	64.29	44.1-81.4	60.0	64.3	1.68	0.62
66	56.00	34.9-75.6	64.29	44.1-81.4	58.3	62.1	1.57	0.68
68*	56.00	34.9-75.6	71.43	51.3-86.8	63.6	64.5	1.96	0.62
70	52.00	31.3 - 72.2	71.43	51.3 - 86.8	61.9	62.5	1.82	0.67
71	44.00	24.4-65.1	78.57	59.0-91.7	64.7	61.1	2.05	0.71
73	32.00	14.9-53.5	78.57	59.0-91.7	57.1	56.4	1.49	0.87
74	24.00	9.4-45.1	85.71	67.3-96.0	60	55.8	1.68	0.89
76	24.00	9.4-45.1	92.86	76.5-99.1	75.0	57.8	3.36	0.82

Note: Absent cut scores reflect ties. CI = Confidence Interval; PPV = Positive Predictive Value; NPV = Negative Predictive Value; LR+ = Positive Likelihood Ratio; LR- = Negative Likelihood Ratio.
* Highest Youden index

analysis, in general, positive likelihood ratios of one or greater are preferable and can be found on the Parent ASRS with a cut score of 73 and below and at all scores for the Teacher ASRS. The greatest positive likelihood ratio for the Parent ASRS is 73 and for the Teacher ASRS is 76. The most optimum negative likelihood ratio falls at 61 for the Parent ASRS and 68 for the Teacher ASRS. Predictive power is maximized at 69 for the Parent ASRS and 76 for the Teacher ASRS. In general, results of the utility analysis indicate that scores lower than the recommended cut score (T-score of $\geq$ 70) may be better to consider when attempting

to differentiate children with a potential ASD or DD in a population of preschool-age children referred for special education eligibility determination.

Research Question 3

The third research question sought to determine the AUC of the optimal cutoff score of the ASRS that optimizes sensitivity and specificity when used with a population of children referred for special education eligibility determination. ROC analysis examined the ASRS's sensitivity to the instrument's false positive rate. The ROC plots the false positive rate on the x-axis (or 1 – specificity) and the sensitivity on the y-axis; the result is the contrast between the two rates. The closer the AUC is to 1, the better the discriminant ability of the test. Figure 9.3 demonstrates the ROC curve for the Parent, Teacher, and Short Forms of the ASRS.

In the current study, the AUC statistic for the Parent ASRS was .58, for the Teacher ASRS was .62, and for the Short Form of the ASRS was .64. The cut score to optimize the sensitivity and specificity of the Parent ASRS is a T-score of 69 on the Total Scale. Sensitivity was 64.86 (95% CI [47.5, 78.9]), and specificity was 63.30 (95% CI [45.5, 79.0). For the Teacher ASRS, the cut score to optimize the sensitivity and specificity is a T-score on the Total Scale of 68. At that score on the Teacher ASRS, sensitivity was 56.00 (95% CI [34.9, 75.6]), and specificity was 71.43 (95% CI [51.3, 86.8]). For the Short Form ASRS, the optimal cut score was 64, resulting in sensitivity of 77.42 (95% CI [58.9, 90.4], and specificity of 48.15 (95% CI [28.7, 68.1].

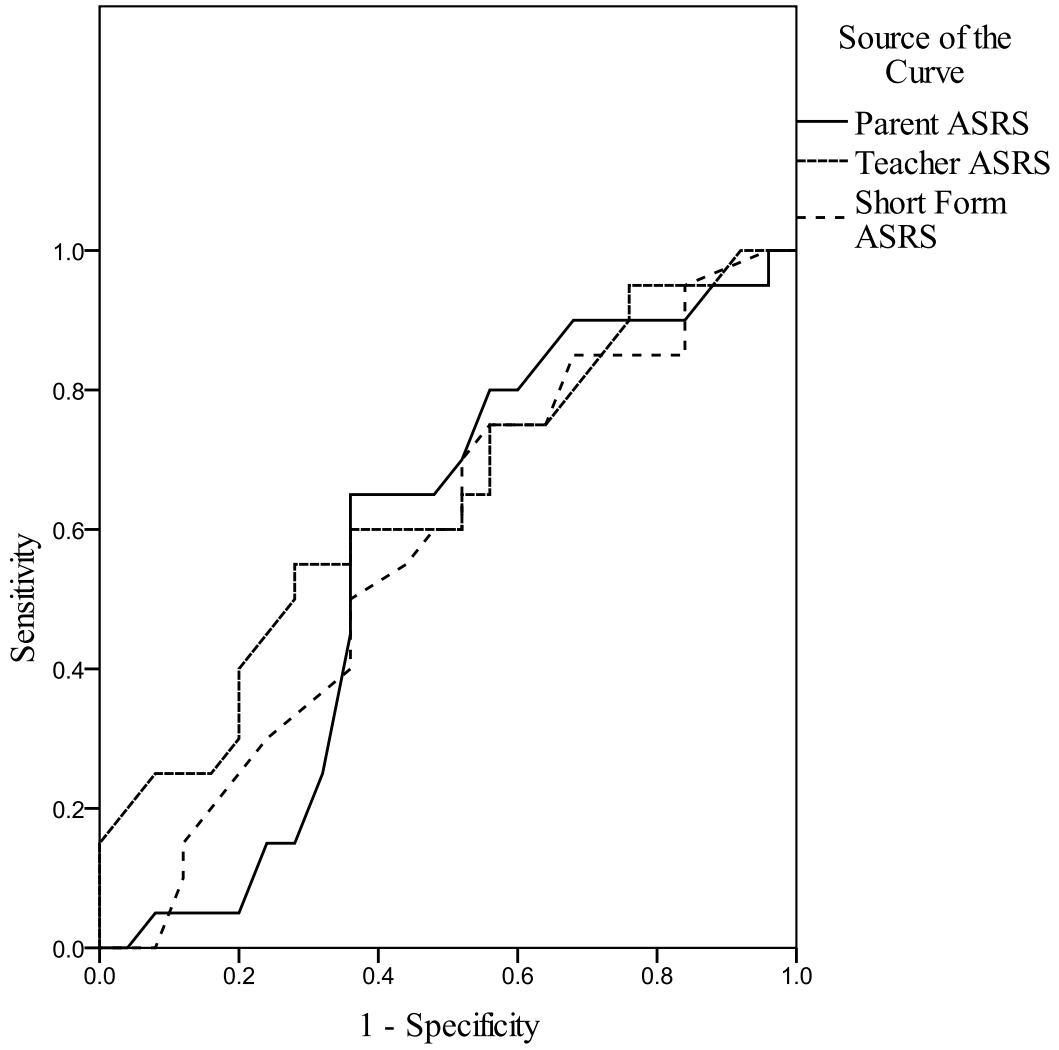

Figure 9.3 *ROC Curve Analysis Mapping Parent, Teacher, and Short Form Versions of the ASRS*

Research Question 4

The fourth research question examined how positive predictive value and negative predictive value, and possible changes in cut scores would be expected to change with different base rates that might be found in different settings. A major factor influencing the utility and Type I and Type II error of any diagnostic test is the prevalence of the disorder the test may be examining. When considering the different aspects of clinical utility, sensitivity,

and specificity, LR+ and LR- are not affected by the prevalence of a disorder, but predictive values are (Ioannidis & Tatsioni, 2010). Thus, predictive values can be leveraged to examine the effects the prevalence of a disorder may have on the effectiveness of a test in different contexts. Given comparable findings of the Parent and Teacher versions of the ASRS to the Short Form Version of the ASRS, analyses to address research question 4 was performed only on the full length versions of the Parent and Teacher ASRS.

As mentioned, predictive values can be used to examine how the ASRS is likely to perform in different contexts where the prevalence of ASDs is likely to differ. Within a special education setting the prevalence of children with ASDs has been reported to be approximately 6.8% (U.S. Department of Education, 2010). At the recommended cut score, the positive predictive value and negative predictive value for the Parent ASRS was 64.86 and 63.30, respectively. For the Teacher ASRS, the positive predictive value and negative predictive value was 52.00 and 71.43, respectively. The effectiveness of the ASRS in differentiating between groups of children with an ASD and a general DD in settings with different base rates can be seen by comparing the sensitivity and specificity of the instrument with the predictive values.

To examine the difference in performance of the ASRS in the special education population with a different base rate, we can compute the positive predictive value and negative predictive value relative to that particular population. To do this, it is best to start with the knowledge that the prevalence of ASDs in the special education population is 6.8%. Thus, in a sample of 1000 students in a special education setting it would be expected about 68 children would have an ASD. Utilizing the current study's findings of sensitivity and specificity, 64.86 (Sn) and 63.30 (Sp) for the Parent ASRS and 52.00 (Sn) and 71.43 (Sp) for

the Teacher ASRS, we can calculate the predictive power for the special education prevalence rate. In a special education population, out of 1000 children, with a base rate of 6.8%, 68 children would be expected to have an ASD. At this prevalence rate the Parent ASRS would identify 44 as having an ASD, with 24 false negatives. The Teacher ASRS would identify 35 of the children as having an ASD, with 33 false negatives. Of the 932 children with no ASD, the Parent ASRS would correctly classify 590 as not having an ASD, with 342 false positives. For the Teacher ASRS, 665 would be identified as not having an ASD, with 267 false positives.

Placing these numbers in the formulas for predictive values, for the Parent ASRS, positive predictive value = PPV = [44/(44+342) = .11 and negative predictive value = NPV = [590/(24+590)] = .96, and for the Teacher ASRS, PPV = [(35/(35+267) = .12 and NPV = [665/(33+665)] = .95 . This means that in the special education population a positive outcome on the ASRS has an 11% chance of being correct for the Parent ASRS and a 12% chance of being correct for the Teacher ASRS. A negative outcome has a 96% chance of being correct for the Parent ASRS and a 95% chance of being correct for the Teacher ASRS.

Taken together, in the base rate condition of the special education context, the ASRS has a more difficult time detecting an ASD than in the study sample (with a prevalence of roughly 55%); however, the results of the prospective analysis, taking into account the base rates of the children with ASDs in the special education population, is likely to be more relevant to the context in which the ASRS will be used. Thus, the relevance of the current scores may address how the ASRS may perform for a practitioner in the special education context. However, as the context and prevalence change the ASRS is likely to perform differently, as the classification statistics of any research are not static.

Supplemental Research Question 1

Supplemental Research Question 1 addressed if differences in ASRS cut scores would be beneficial when considering an individual's IQ, language ability, and adaptive skills. As before, given the similarities between the psychometric properties of the different forms of the ASRS found in this study, only the Parent ASRS was utilized in this analysis. When examining the effects of different cut scores on the areas of IQ, adaptive and language abilities, scatter plots were used to illustrate how participants scored across different areas of development and how cut scores may alter the Parent ASRS hit and error rates. Figures 10.3, 11.3, and 12.3 indicate where individuals scored in the areas of intelligence, adaptive abilities and language, respectively. Each figure plots those with standard scores above 70 with light circles and those at 70 or below with dark diamonds, dotted lines separate the matrix into the classification and error cells. Higher scores on each measure indicate more characteristics of an ASD. By moving the vertical axis of each scatter plot to represent different cut scores insight develops into what forms of error become more prevalent within each domain.

With a cut score at the clinically significant level (T-score $\geq$ 70) on the ASRS those individuals with IQ scores two standard deviations (standard score $\leq$ 70) below the mean (standard score mean = 100; standard deviation = 15) or lower had a hit rate of 71%; Type I error rate was 10% and Type II error rate was 19%. For those children within two standard deviations of the mean (standard score of $\geq$ 71) there was a 61% hit rate; Type I error rate was 20% and Type II error rate was 18%. In the area of Adaptive functioning scores two standard deviations (standard score $\leq$ 70) below the mean (standard score mean = 100; standard deviation = 15) or lower had a 68% hit rate; Type I error rate was 8% and Type II error rate was 24%.

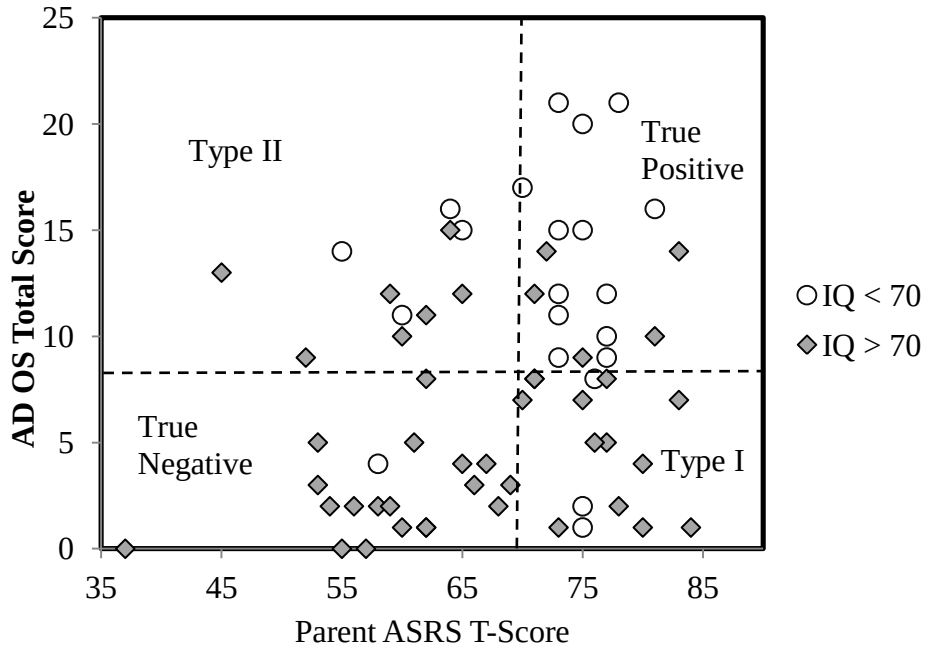

Figure 10.3 *Scatter Plot Illustrating Classification and Error of the Parent ASRS Based on IQ*

.

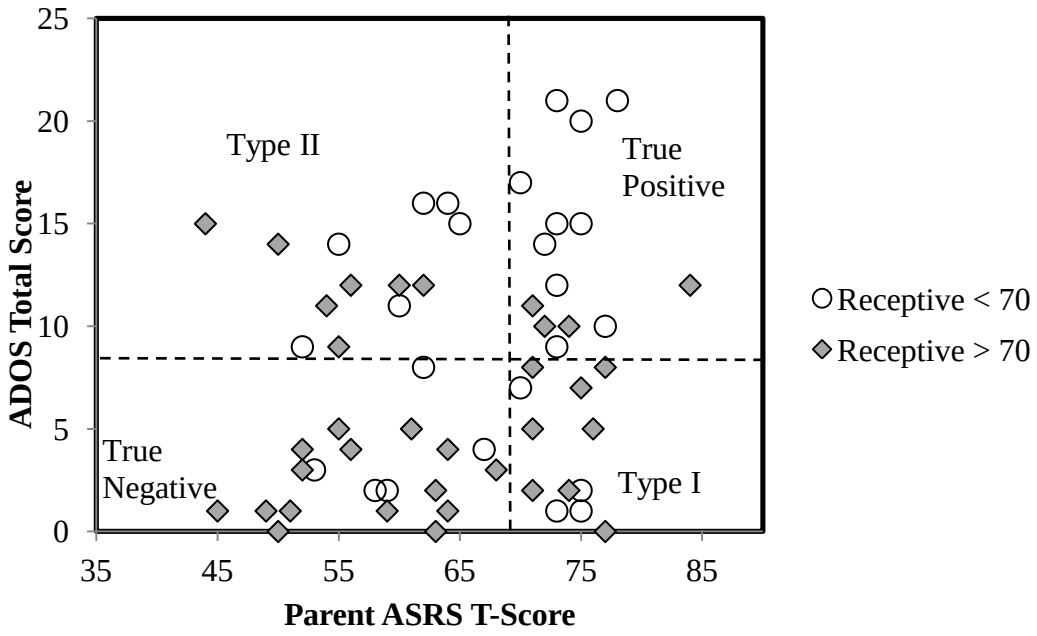

Figure 11.3 *Scatter Plot Illustrating Classification and Error of the Parent ASRS Based on Receptive Language*

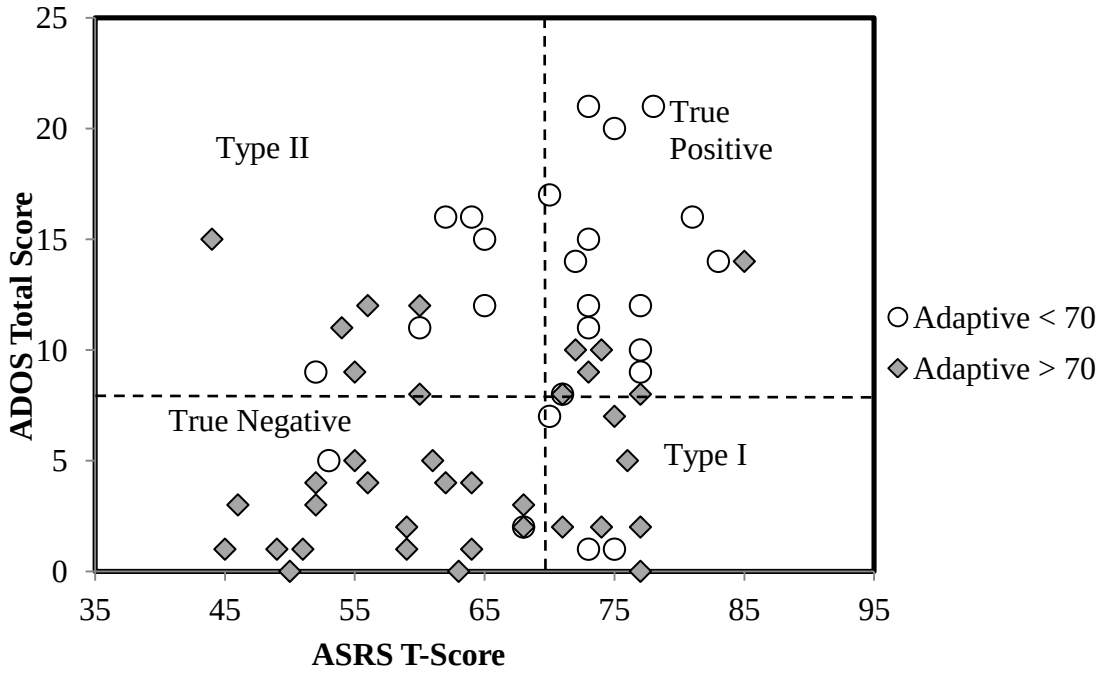

Figure 12.3 *Scatter Plot Illustrating Classification and Error of the Parent ASRS Based on Adaptive Skills*

For those children within two standard deviations of the mean (standard score of ≥ 71) there was a 60% hit rate; Type I error rate was 23% and Type II error rate was 18%. For Receptive Language, scores two standard deviations (standard score ≤ 70) below the mean (standard score mean = 100; standard deviation = 15) or lower had a 60% hit rate; Type I error rate was 16% and Type II error rate was 28%. For those children within two standard deviations of the mean (standard score of ≥ 71) there was a 61% hit rate; Type I error rate was 22% and Type II error rate was 17%. For Expressive Language, scores two standard deviations (standard score ≤ 70) below the mean (standard score mean = 100; standard deviation = 15) or lower had a 57% hit rate; Type I error rate was 17% and Type II error rate was 27%. For those children within two standard deviations of the mean (standard score of ≥

71), there was a 65% hit rate; Type I error rate was 19% and Type II error rate was 16%. Table 13.3 outlines the hit and error rates for each of the developmental areas. Figures 11.3 and 12.3 demonstrate scatter plots of language and adaptive skills. Visual inspection of the vertical axis on the developmental scatter plots shows that lowering the cut score changes the amount of Type I error for those children two standard deviations below the mean or lower, and a selected set point can be established based on what error types are more acceptable. Based on the Parent ASRS cut score of 70 each developmental area can be explored in terms of the types of hit rates and errors that may result. When considering intelligence, adaptive abilities, and expressive and receptive language, there are more Type I errors for those who performed across the developmental areas with standard scores $\geq$ 70. The Type I error rate of children with standard scores across all developmental variables of $\leq$ 70 was found to be at 29%, whereas the Type I error rate for those children whose standard scores on the developmental variables were $\geq$ 70 was 71%, a 42% difference. Figure 13.3 illustrates the percentage of classification type and error rates by the combined developmental areas under two standard deviations from the mean.

The dark bars represent children with combined scores of IQ, adaptive skills, receptive and expressive language abilities all two standard deviations or below (standard score $\leq$ 70). Light gray bars represent children with combined scores of IQ (i.e., Leiter or Stanford-Binet – Fifth Edition), adaptive skills (SIB-R), receptive (i.e., PPVT-4) and expressive (i.e., EVT-2) language scores within two standard deviations of the mean (standard score $\geq$ 71). The graph is separated into true classifications and error rates for each developmental area. Each error and classification rate is listed by the percentage that they are meeting for each respective area.

Table 13.3: *Summary of Classification and Error Rate by Developmental Ability Level on the Parent ASRS*

Developmental Area		Error Type			
	True Positive	Type I Error	Type II Error	True Negative	Hit Rate
IQ ≤70	5	10	19	67	71
IQ > 70	41	20	18	20	61
Adaptive ≤ 70	8	8	24	60	68
Adaptive > 70	43	23	18	18	60
Receptive ≤ 70	12	16	28	44	60
Receptive > 70	39	22	17	22	61
Expressive ≤ 70	17	17	27	40	57
Expressive > 70	42	19	16	23	65
All Areas ≤ 70 Total (%)	15	29	49	63	*
All Areas > 70 Total (%)	85	71	51	37	*

Note: Table is presented in percentages. IQ = Stanford-Binet – Fifth Edition or Leiter-R; Adaptive = Scales of Independent Behavior – Revised; Receptive = PPVT-4 = Peabody Picture Vocabulary Test – Fourth Edition; Expressive = EVT-2 = Expressive Vocabulary Test – Second Edition.
* Not Applicable.

Taking into consideration those children with developmental variables with standard scores ≤ 70 and those > 70, Type II error was about even. In short, for the areas of adaptive skills, receptive language, and expressive language, a greater number of true positives are found when standard scores are 70 or above in each of the developmental areas

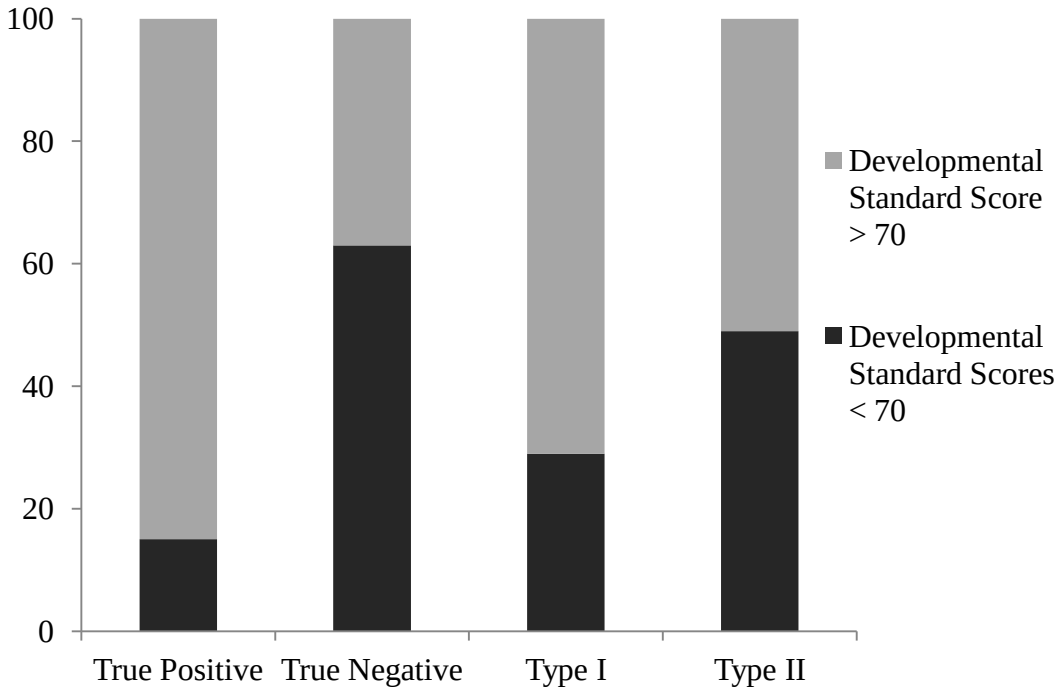

Figure 13.3 *Illustration of ASRS Classification and Error Type by Developmental Areas*

Supplemental Research Question 2

The final set of data analyses sought to determine which specific items from the ASRS discriminate most between children on the autism spectrum and children with other developmental disabilities. A point-biserial correlation was performed using the items from the Parent ASRS in order to determine each item's ability to discriminate between children in either the ASD or DD groups. Figure 14.3 lists the point-biserial correlations for each item of the Parent ASRS. The ASD group was listed as 1 while the DD group was listed as 0 for the binary grouping variable. Higher scores on the ASRS items indicate a greater likelihood for behaviors typical of children on the autism spectrum. The scaling of the items was as follows: 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Frequently, 4 = Very Frequently. Items

were reverse-scored for the analysis as necessary. For the Parent ASRS, the top five item descriptors included: "Point to objects when asked to," "Keep a conversation going," "Fail to make his/her needs known," "Use make believe play," and "Understand the point of view of others." The ASRS demonstrated similarities to previous research findings indicating that items on measures with the greatest ability to differentiate between groups of children with a potential ASD and other types of DDs reflect socialization and social interaction factors (Ventola et al., 2007). Table 14.3 and 15.3 outline the most discriminating items from the Parent and Teacher ASRS, respectively.

The individual items on the Parent ASRS that were found in the current study to be the most discriminating between groups differed from the Short Form ASRS items. The 15 items that make up the Short Form ASRS, taken from the long version of the Parent ASRS, are items 4, 5, 15, 24, 29, 42, 43, 44, 46, 47, 49, 50, 51, 52, and 61. The top 15 discriminating items of the Parent ASRS were 7, 29, 67, 18, 14, 39, 40, 66, 28, 44, 49, 51, 21, 15, and 50. Only 33% of the most discriminating items found in the current study on the Parent ASRS and the Short Form ASRS were the same on both scales. It should be noted that there were also a number of negative correlations, a less-desired outcome when attempting to discriminate between the ASD and DD groups. This indicates that the scoring is doing the opposite of how the measure may have intended to use the items for discriminating between groups. Items that correlated with a coefficient at r = -.20 or greater in magnitude when numbered in rank from greatest to least were 6, 41, 26, 58, 31, and 17.

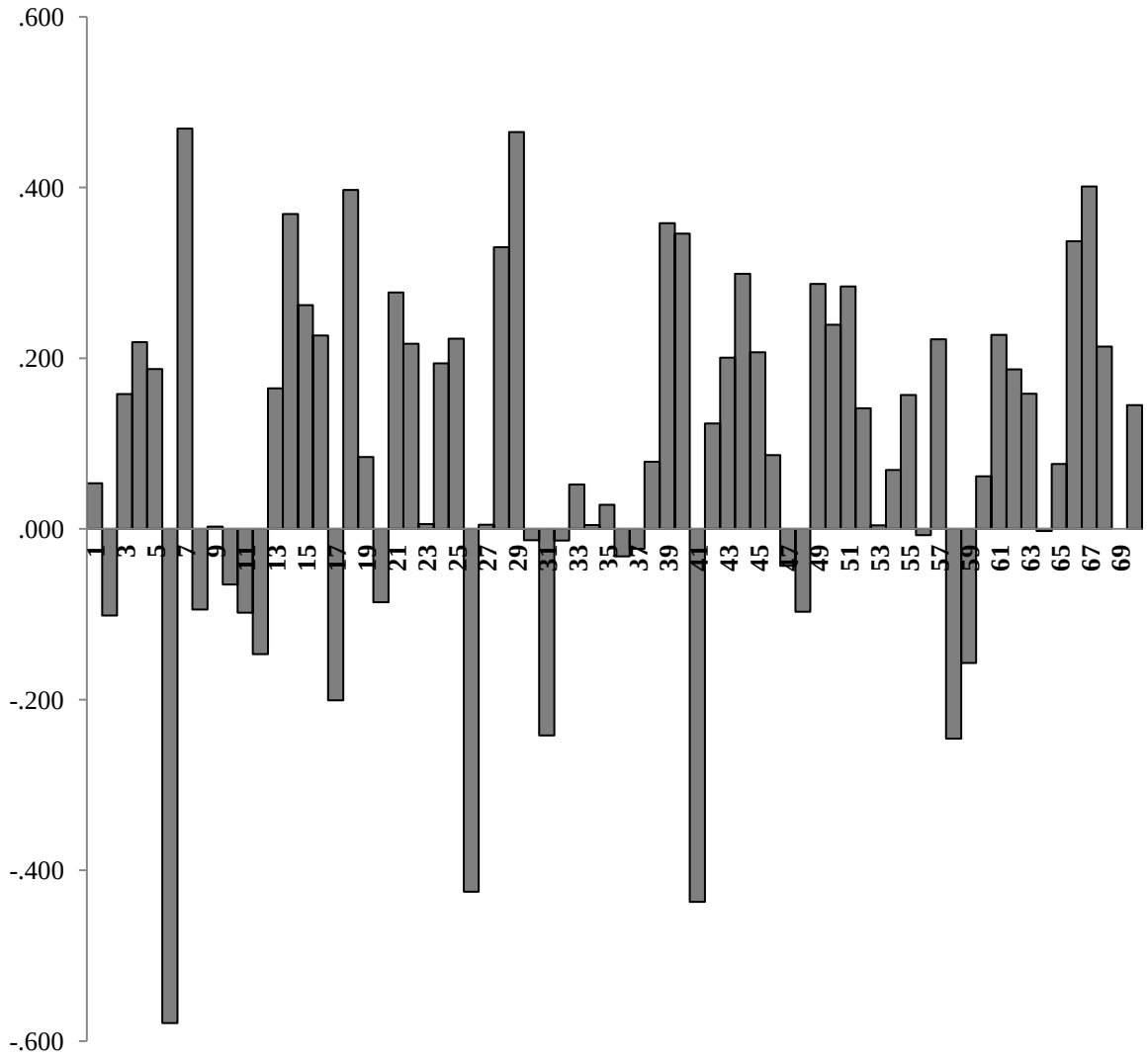

Figure 14.3 *Item Point-biserial Correlations of the Parent ASRS*

Table 12.3: *Most Discriminating Parent ASRS Items*

Item	Descriptor	*r*	ASD Mean	DD Mean
7	Point to objects	0.47***	2.03	0.89
29	Keep conversation	0.47***	3.06	2.00
67	Fail to make needs known	0.40**	2.35	1.19
18	Use make believe	0.40**	1.97	1.04
14	Understand others	0.37**	2.97	2.19
39	Fascinated with object parts	0.36**	3.06	2.22
40	Respond to other children	0.35**	1.97	1.30
66	Smell, taste, eat inedibles	0.34*	1.04	0.31
28	Start conversation	0.33*	2.59	1.74
44	Trouble talking with adults	0.30*	2.65	1.89
49	Seek company of children	0.29*	2.34	1.70
51	Social problems with children	0.28*	2.69	2.11
21	Respond to adults	0.28*	1.75	1.19
15	Talk to other children	0.26*	2.94	2.30

Note: ASD = Autism Spectrum Disorder; DD = Developmental Disability
* $p < .05$
** $p < .01$
*** $p < .001$

These items included, “Ask questions off topic” (ASD mean = .59; DD mean = 2.07), “Talk too much to adults” (ASD mean = .72; DD mean = 1.85), “Talk too much to children” (ASD mean = .66; DD mean = 1.70), “Interrupt others” (ASD mean = 2.16; DD mean = 2.74), “Get into trouble” (ASD mean = 1.56; DD mean = 2.07), and “Disorganized” (ASD mean = 1.28; DD mean = 1.70). On the Teacher ASRS, negative correlations showed up in the biserial correlations, as well. Items with correlation coefficients at $r = -.20$ or greater in magnitude in order from greatest to least were 6, 26, 8, 11, 58, and 60. These items included: “Ask questions off topic” (ASD mean = 0.79; DD mean = 1.77), “Talk too much to children” (ASD mean = 0.50; DD mean = 1.35), “Insist on same way” (ASD mean = 1.38; DD mean = 2.12), “Line up objects” (ASD mean = 0.88; DD mean = 1.54), “Interrupt others” (ASD

Table 15.3: *Most Discriminating Teacher ASRS Items*

Item	Descriptor	R	ASD Mean	DD Mean
29	Keep conversation	0.55***	3.25	1.96
7	Point to objects	0.50***	2.04	0.88
14	Understand others	0.43***	3.13	2.31
3+	Understand others	0.41**	2.92	2.12
67	Fail to make needs known	0.40**	2.35	1.19
15	Talk to other children	0.38**	3.13	2.19
52+	Understand humor	0.38**	3.00	2.12
53+	Repeat words	0.38**	1.67	0.69
28	Start conversation	0.38**	2.92	2.04
66	Smell, taste, eat inedibles	0.34*	1.04	0.31
22+	Immature Language	0.34*	2.63	1.62
4+	Play with peers	0.33*	2.42	1.58
43+	Avoid looking at people	0.33*	2.46	1.73

Note: ASD = Autism Spectrum Disorder; DD = Developmental Disability;
+Items unique to the Teacher ASRS when compared to top discriminating items of the Parent ASRS
* $p < .05$
** $p < .01$
*** $p < .001$

mean = 1.63; DD mean = 2.15), and "Detail-obsessed" (ASD mean = 0.79; DD mean = 1.27). The Teacher ASRS items with negative correlations of $r \leq$ -.20 overlapped 50% with the Parent ASRS items. Appendix H outlines the biserial correlations and means for the ASD and DD groups for each of the items of the Parent ASRS. Appendix I outlines the biserial correlations and means for the ASD and DD groups for each of the items of the Teacher ASRS.

CHAPTER IV

DISCUSSION

The rates of children with an autism spectrum disorder (ASD) are increasing and prevalence rates are on the rise in both the general population and special education contexts. Early identification and intervention efforts of children with ASDs have benefits that extend into later life, facilitating greater long-term successes. A child's preschool years are a time when parents and professionals typically first identify the symptoms of ASDs. As a result, efforts have been made to improve the early identification and intervention of preschool children with ASDs. Although federal funding for ASDs has increased over the past several years, the public schools have not seen much of this monetary benefit. As a result, resources have become limited in the schools making the comprehensive and accurate identification of children with ASDs increasingly difficult. It is essential that the public school system be able to efficiently identify children who may be on the autism spectrum in order to ensure access to early intervention services.

Public school budgets continue to be cut and schools are increasingly left short-handed in terms of personnel who have expertise in autism. Moreover, resources to provide comprehensive assessments of children suspected to have disabilities also become limited as public education funds become scarce. As a result, school professionals are becoming more dependent on parent and teacher behavior checklists as a primary means of identifying

children with ASDs, a practice that goes against best practice guidelines. To complicate matters further, the checklists being used often do not have sound diagnostic utility and may subsequently over- or under-identify children with ASDs. Research has demonstrated that only a limited number of diagnostic assessment tools can reliably identify children with ASDs. Unfortunately, most of the measures that can successfully identify children with ASDs involve intensive training and time that is not widely available due to limited resources in the public schools. There is a need for a psychometrically-sound behavior checklist that can be used to reliably identify children who may have an ASD in the public schools.

The current study examined several utility indices of a new behavior checklist, the Autism Spectrum Rating Scales (ASRS), such as sensitivity, specificity, predictive values, and likelihood ratios, to determine the measure's ability to identify and differentiate between children on the autism spectrum from other children with a general DD in a population of preschoolers referred for special education services. Additionally, the study set out to explore how the ASRS would perform in contexts with attention to specific base rates and examined the acceptability of different classification and error rates in the different contexts in which the ASRS may be used.

Classification of the ASRS

The recommended cut score for the current study was based on the general rule of thumb for "clinical significance." When based on a T-score this generally starts at the 98th percentile (T-score $\geq$ 70). In the current study, at a T-score of $\geq$ 70, the Parent ASRS (Long Form) was able to correctly classify 64% of the children either as having an ASD or having a general DD. For the Parent ASRS, false positives were found at a rate of 16%, whereas false

negatives were found at a rate of 19%. The Teacher ASRS (Long Form) was able to correctly classify 62% of the children either as either having an ASD or a general DD. Additionally, for the Teacher ASRS, 15% of the children identified as having an ASD were false positives and 23% who were not identified were false negatives. Overall, the classification abilities of the Parent and Teacher ASRS were relatively similar. Thus, it appears that both forms of the ASRS, Parent and Teacher, appear to be fairly comparable in terms of their ability to classify children with ASD.

The Short Form of the ASRS was comparable to the long versions of the Parent and Teacher ASRS, indicating that the utility of the Short Form ASRS may be similar in nature to the longer versions. The overall hit rate of the Short Form ASRS was 62% with false positives at 17% and false negatives at 21%. In general, these properties are close to those found on the Parent and Teacher ASRS, indicating that the Short Form ASRS may perform similarly to the longer ASRS versions for the 2-to 5-year-olds. Using the Short Form ASRS, therefore, may be beneficial for use in situations that have particular time and resource constraints. Across all forms (i.e., Parent ASRS, Teacher ASRS, Short Form ASRS) hit rates and false positive rates were within 2 percentage points of each other and false negatives were within 4 percentage points of each other, demonstrating a general consistency in performance for each form.

Further analysis allowed insight into the utility of the ASRS when used with a preschool population referred for possible special education services. Using the recommended cut score, positive predictive values indicated that on the Parent ASRS, 69% of the individuals with a positive result on the ASRS actually had ASD, whereas on the Teacher ASRS the same was true for 62%. Thus, a positive result on the Parent ASRS may

hold only slightly more weight diagnostically when attempting to rule in a positive identification of a child with a score at or above the cutoff threshold for an ASD, although there was only a difference of 8 percentage points between ASRS forms. In contrast, 59% of children were identified by the Parent ASRS as having another type of DD, and 63% of children having another form of a DD were identified by the Teacher ASRS. Again, both forms appear to be comparable in their ability to determine if a student does not have an ASD, with a difference of only 4 percentage points between the negative predictive values at the recommended cut score.

The results of the ROC Curve analysis determined the AUC values at the optimal cut scores of the ASRS, where sensitivity and specificity are optimized. ROC Curve analysis indicated that the AUC was .58 at the optimal cut score for the Parent ASRS (T-score of 69). For the Teacher ASRS, the cut score for optimal sensitivity and specificity was 68 where the AUC was .62. For the Short Form ASRS, the optimal cut score was 64 where the AUC was .64. Similar to previous analysis, the AUC statistics were approximately the same between forms with a difference of only .06 points between all the forms. Thus, at their optimal sensitivity and specificity, all ASRS forms tend to perform similarly. It should be noted, however, that the ROC Curve analysis only allows insight into the best cut score for the instrument's optimal sensitivity and specificity and does not directly address the instrument's performance and general clinical utility (e.g., classification and error rates) in different types of settings and conditions. Other indicators of the measure's utility such as the predictive values described above can provide more insight into the instrument's clinical performance.

Table 16.4 outlines the performance of the ASRS compared with other commonly used measures to assess for ASDs. Because each measure has its own set of studies from the

respective standardization samples to determine how well the measure classifies children into ASD and non-ASD groups, each based in different sample characteristics, efforts were made to create continuity in comparing the data. The average figures for sensitivity and specificity as reported by the publisher's standardization studies are reported in Table 16.4 and were used to find the positive predictive values and negative predictive values as they pertain to the population and prevalence rates of the current study. Therefore, a degree of standardization among measures could be accomplished for better ease of comparison.

Table 16.4: *Comparison of ASRS with Other Current Measures Standardized to Current Sample*

Measure	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**Type I**	**Type II**	**Hit Rate**
Parent ASRS[a] Current Study	64.86	63.30	.69	.54	16%	19%	64%
ASRS[a] Standardization Sample	92.6	91.55	.92	.90	4%	4%	91%
CARS-2[b]	.86	.73	.80	.90	12%	7%	81%
GADS[c]	.86	.70	.78	.81	13%	7%	79%
GARS-2[d]	.92	.85	.89	.90	6%	4%	90%

Note: Figures of sensitivity and specificity are derived from the average of the reported studies in the standardization samples for each respective measure. PPV = Positive Predictive Value; NPV = Negative Predictive Value; Type I = Type I Error (false positive); Type II = Type II error (false negative); ASRS= Autism Spectrum Rating Scales; CARS = Childhood Autism Rating Scale; CARS-2; Childhood Autism Rating Scale – Second Edition; GADS = Gilliam Asperger's Disorder Scale; GARS-2 = Gilliam Autism Rating Scale – Second Edition.
[a] Goldstein & Naglieri, 2010; [b] Schopler, Van Bourgondien, Wellman, & Love, 2010
[c] Gilliam, 2003; [d]Gilliam, 2006.

The results of the current study are also included in Table 16.4. Overall, the ASRS has psychometric properties that may make it a valid instrument in the early identification of children with possible ASDs. It should be noted that even the best measure will have difficulty differentiating between groups of individuals with similar overt behavioral presentations (i.e., shared variance), particularly when the disorder the instrument is attempting to identify has a low base rate.

Determining the Optimal Cut Score on the ASRS

In exploring the possibility of using different cut scores of the ASRS, based on study results, the current research indicated that different cut scores may be beneficial when working with a more specialized population such as a preschool special education population. According to the indices of utility (e.g., sensitivity, predictive values) derived from the results of the current study, the findings indicated that lower cut scores may be better used in the identification of children with potential ASDs because of better statistical properties. It is felt that the classification errors that may be made in the process of changing the cut score are tolerable given the circumstances in which the measure is being used. For example, when using the ASRS with a population of preschool children referred for special education services it may be beneficial to utilize a lower T-score in the identification of children with a potential ASD to reduce the risk of missing children at an age when a diagnosis is often more difficult to make and the error in a false positive (e.g., has access to services) has less severe consequence than a false negative (e.g., denied services).

One issue complicating the decision of what cut score to utilize with a preschool population referred for potential special education services would be the fact that preschool

children with ASDs often share behavioral similarities, symptom presentation and comorbidities with other diagnoses from which the ASRS may be attempting to differentiate (Fernell et al., 2010). Because of the shared variance, or similarity in behavioral presentation between groups, the ASRS, as well as other measures that seek to differentiate autism from other forms of developmental disabilities, will inherently have a more difficult time accomplishing this task.

As touched on briefly, the needs of the setting as well as the populations of children that the instrument is being used with are critical factors to consider when determining if a different cut score will better identify children on the autism spectrum. In a setting where there is more symptom overlap among populations, such as in the case of the current study where the ASRS attempted to differentiate between children with an ASD from children with a general DD, utilizing a lower cut score may allow the measure to better distinguish between the groups. In contrast, in a setting where the intent is to distinguish between children with a possible ASD and another group of children that does not share much symptom overlap (such as in a pediatrician's office), a higher cut score may be more relevant and useful.

Modifying Cut-Scores

Environmental context affecting cut scores. The results of the positive and negative predictive values obtained in the current study show how different base rates would change the utility of the ASRS depending on the prevalence rates of ASDs. In the special education population, where the prevalence rates of an ASD are about 6.8%, the positive predictive value for the Parent ASRS was .11 and the negative predictive value was .96. For the Teacher ASRS, within the special education population, the positive predictive power was

.12 and the negative predictive power .95. Taken together, in the special education context, negative results are more likely to be correct given the low base rate of ASD in the population. To illustrate, with an ASD base rate of 6.8%, if a clinician were to always state that a child did not have an autism spectrum disorder, the statement would be correct about 93% of the time. With the ASRS, however, the clinician can do better than always failing to diagnose ASD about 2-3% of the time. By always diagnosing an autism spectrum disorder, a clinician would be correct about 7% of the time. Thus, using the measure increases the confirmation of an ASD about 4-5%. In all, there are limited benefits to using the ASRS when compared to never making the diagnosis of ASD. However, when applied in a setting where long-term successes and benefits in life depend on appropriate identification of an ASD (e.g., early identification of children on the autism spectrum), any gain in diagnostic prowess, no matter how small, may be welcome.

Child characteristics affecting cut scores. Although only viewed as exploratory in nature due to small sample size, when examining if there may be differences in ASRS cut scores based on a child's IQ, language ability, and adaptive skills, the current study found that hit and error rates tended to be contingent upon certain developmental characteristics of a child. Classification and error rates for the different developmental variables were examined by looking at those children who were performing at a level of two standard deviations below the mean (standard score $\leq$ 70) and those within two standard deviations (standard score > 70) of the mean (mean = 100, standard deviation = 15). On tests of IQ, adaptive and language abilities, a greater percentage of true positives was found with children considered "higher-functioning;" that is, participants who had standard scores of 70 or above on measures of IQ, adaptive abilities, receptive language, and expressive language

tests when compared to those who scored at or below a 70, or, "lower-functioning" children on the same measures. This phenomenon suggests that there is a greater chance that the ASRS is able to correctly identify ASDs in children who perform within two standard deviations of the mean, or those with standard scores greater than 70 on tests such as the Stanford-Binet – Fifth Edition and the Scales of Independent Behavior – Revised. However, Type I error also tended to be found in greater percentages among higher-functioning children, meaning that there is a greater chance of the ASRS making a false positive error with a higher-functioning student when compared to a lower-functioning student. With lower-functioning children, or those who had standard scores $\leq$ 70 in the areas of IQ, adaptive skills, receptive language and expressive language, the ASRS tended to be better at identifying children who had a general DD. That is, higher percentages of true negatives was found by the ASRS among those children who scored lower on each of the developmental variables assessed in the study. Type II error tended to be relatively equal for the two developmental levels, indicating the ASRS has about the same chance of making a false negative error regardless of how the child is functioning across a number of developmental areas.

The findings of the study have implications for clinicians using the ASRS in a diagnostic context. Findings indicate that if a student has scores in several major areas of development that are above a standard score of 70, and there is a positive identification for an ASD, the positive identification may be more accurate than for a student who has scores below a standard score of 70 across the same areas of development. To be identified as not having an ASD will be more meaningful for those children with scores across several major areas of development at or below a standard score of 70 on the ASRS, but less so for the

children who have standard scores above 70 in the same developmental areas. More accurate identification of children with standard scores on developmental variables of under 70 can be made by lowering the cut score, whereas with those individuals with standard scores above 70 on the developmental variables, increasing the cut scores will better help correctly identify children on the autism spectrum.

Assessment of ASRS Test Items

Previous research has shown that test items pertaining to socialization and social interaction have been best at discriminating between groups of children with an ASD and other groups of children with similar developmental disabilities (Ventola et al., 2007). Similar to previous research, the current study found that test items on the ASRS that tended to best discriminate between the ASD and DD groups reflected variables relating to social interactions. The most discriminating items were those that pertained to verbal communication (e.g., initiating and maintain a conversation, making needs known, and communicating nonverbally by pointing). Social interaction variables were also represented in items that were most highly discriminating between the ASD and DD groups. Among items that reflected social interactions included those that inquired about the child's ability to understand other people's communications and intentions, responses to other children, and efforts to be with other children. Interestingly, the ASD group scored better on some of these items when compared to the DD group which suggests more social interest, interaction, and communication. Given the fact that the DD group also had communication delays, this finding may not be so surprising, and again, point to why it is that professionals find it so difficult to distinguish children with ASDs and other developmental disabilities when they

are preschool age. Fortunately, there were other ASRS items that discriminated between the ASD and DD groups. Among those items were those that pertained to stereotyped and sensory behaviors (e.g., fascination with parts of objects and sniffing and tasting inedible objects). This may indicate that the overt presentation and most discriminating symptoms and behaviors of children with ASD are not limited to socialization and social interaction factors alone, but to a more broad range of items that are reflective of other symptoms displayed by children with ASDs as well.

Of note is the fact that when the point-biserial correlation analysis was performed on the ASRS items, there were a number of items that were negatively correlated. The negative correlations may indicate that the item is performing in a manner that is the opposite of how the item was originally intended. A number of these items had to do with verbal communication (e.g., "Ask questions off topic," "Talk too much to adults," "Talk too much to children"). When examining these items, the means for the ASD group were lower than the DD group, meaning the items are more reflective of the ASD group. Thus, it appears that when examining the items that most discriminate between groups, various verbal communication items appear to be discriminating in opposite ways. For example, the ASD group scored higher than the DD group on the verbal item, "Start conversation," but lower than the DD group on the item, "Talk too much to children."

Context-Specific Decisions and the ASRS

The ASRS can provide important diagnostic information. Whether to use this scale, and rely on the recommended cut-score or modify it somewhat, needs to be determined by the context of the assessment and the costs of misclassification. Decision theory proposes

that perceived costs of misclassification within any context be examined in terms of monetary costs, as well as effects on individuals (Winkler & Hayes, 1975).

Within the architecture of decision theory, multiple perspectives should be examined when determining costs of misdiagnosis, including those of respondents as well as the institutions with which respondents may have contact, such as health and research facilities (Smits, Smit, Cuijpers, & De Graaf, 2007). As previously discussed, difficulty with the ASRS as well as with other measures that seek to differentiate children with potential ASDs from children with general DDs may have to do with behavioral and symptom overlap as well as low base rates of ASDs. That is, instruments generally have difficulty identifying a condition when the base rates of the disability are very low or when there is a good deal of shared variance between groups. However, depending on the context in which the ASRS is being used, specific errors may be more or less tolerable in terms of cost to the responder or individual.

In general, a screening setting such as a pediatrician's office or even within a context where children are referred for potential special education services, a false negative on the ASRS may mean the denial of the opportunity for interventions early in life. Early intervention has demonstrated to have beneficial and long-term effects for children on the autism spectrum (Ben Itzchak & Zachor, 2011; Dover & LeCouteur, 2007; Goldstein, Naglieri, & Ozonoff, 2009; Rice, 2007; Rogers, 1996; 1998; Smith, Groen, & Wynn, 2000; Turner, Stone, Pozdol, & Coonrod, 2006). The other form of misclassification, a false positive, will also have costs. In the case of a false positive, costs may come in terms of unnecessary seeking of services and costs to society, and possible deleterious effects to the mental health of the parents who may be involved. In a setting where children are referred for

possible special education services, both types of errors appear to have consequences that have costs to both the institutions and individuals; thus, a clinician must be careful when utilizing the instrument.

The problems with misclassification can also affect researchers who utilize the ASRS to identify children who may be on the autism spectrum, with perceived costs differing for false positive and false negative errors. Thus, researchers may want to assure the most accurate and appropriate identification of children who may be on the autism spectrum. As a result, a false positive may be a more detrimental error than a false negative. The result of a false positive will lead to the wrong group of children being included in studies with subsequent results then being conducted on misclassified groups of children. A false negative may have less impact in a research setting, as a child identified as ASD would then not be included in a study aiming to possibly examine children on the autism spectrum.

In general, a false positive error may be more tolerable in those settings and contexts in which children may be screened for potential early intervention and services. An error in over-identification in these contexts can lead to children receiving early intervention services that may benefit them and their educational and social development later in life. A false negative may be more consequential in settings where children are screened for early intervention, because services may be denied and potential benefits that could have long-term positive effects will be lost. In contrast, in a research or similar setting, a false positive can be a more detrimental error as children outside the intended target population may be included in a study. As previously discussed, this may lead to results that can differ from a more accurately diagnosed sample.

ASRS as Screener Versus Diagnostic Test

The ASRS is an instrument that has the potential to be used as a general screening measure to assess and identify potential children who may be on the autism spectrum. If used primarily as a screener, the ASRS may have disadvantages due to the error rates uncovered in the current study. Ideally, a screener would be able to identify 100% of children who need referrals for more comprehensive diagnostic testing. Operating from the currently recommended cut score (T-score $\geq$ 70), the ASRS would not be able to identify children with potential ASDs at a high rate. Perhaps with some manipulation of the cut score, the ASRS may better serve as a screening measure. However, given the results of the current study it appears that the ASRS may be better used as a diagnostic test that is integrated into a best practice framework (i.e., one used in combination with other assessment measures and diagnostic practices). Along these lines, a clinician can utilize the findings of the current study when using the ASRS as a diagnostic test (e.g., the potential use of higher cut scores with higher-functioning children to avoid Type I error) to ensure that the instrument is performing at its best in identifying children with potential ASDs.

Upon first glance, the Short Form of the ASRS, with its 15 items, may have greater potential than the longer versions of the ASRS to be used as a screener due to its shorter length. However, analyses (e.g., sensitivity, specificity, ROC Curve analysis) in the current study indicated that the Short Form may have a similar hit rate and error profile when compared to the longer versions. Therefore, the Short Form ASRS may not perform to the specifications of an ideal screening measure and capture a high rate of cases, even if some are false positives. However, lowering the cut score may make the Short Form ASRS a more feasible screener as it will be able to identify greater numbers of children who may need

further referrals. Caution, however, needs to be taken when using either the short or long form versions of the ASRS for more general screening versus diagnostic testing if used at the level of the current cut score.

Limitations

Findings should be interpreted within the current methodological context and procedures. It should be noted that there was no control group. The sample of children in the current study all had developmental disabilities and were referred for a special education evaluation. It was the goal of the current study to determine the utility of the ASRS in a setting where children were referred for special education services and its ability to differentiate between children with possible ASDs and DDs in that setting; however, a typically-developing peer sample would have provided different insight into the instrument's performance. Research has demonstrated that there is a great amount of similarity in the behavioral presentations between the groups of children included as the samples in this study at the preschool referral level (Fernell, et al., 2010). Given the extent of symptom overlap and comorbidity between the developmental disabilities of autism and other disabilities included in the current study, such as speech and language impairments, intellectual disabilities, and global developmental delays, it would be reasonably expected that the statistical properties of any measure would be affected by the samples included in the study. The current study examined how the ASRS performed with children referred for potential special education services and the results are reflective of that particular sample, whereas other sampling procedures would likely offer another diagnostic picture of the measure.

An additional consideration when interpreting the results of the current study is the limitations of sample size. In statistical analysis utilizing descriptive techniques, greater sample sizes, theoretically, will be more representative of the true population about which the study is aiming to draw inferences. In the case of the current study, a greater sample may have provided a different picture of the utility of the ASRS that could be potentially more accurate and representative of the greater population of children the ASRS seeks to assess.

Group differences in age also present a limitation. Although efforts were made in the current study to compensate statistically for these differences, having matched groups across the major demographic areas (e.g., age, sex, ethnicity) may offer up a better diagnostic picture of the performance of the ASRS in a situation where early identification and intervention is crucial. It may be important to note that the limitation of age difference between the ASD and DD groups could have been the statistical artifact of a small sample size. Group differences in the ASRS scores could reflect the age differences in the current sample. With a larger sample size, the group difference may have been eliminated. Future research will be needed to determine how age may affect scores on the ASRS.

The grouping technique may also be a limitation to the current study. Any measure has some degree of error. In the current study, the ADOS was used as the criterion measure to group children into either the ASD or DD groups. The ADOS itself commits some classification errors. No criterion measure is able to classify all children without committing some form of Type I or Type II error and the ADOS is no exception. Using a classification method that aligned more with all suggested methods of best practice procedures may have created groups that were more representative of true ASD and DD populations.

Finally, the groups were found to be statistically different in age and there may be variables related to this difference that may affect results. Finding more similarly uniformed groups may provide a different insight into the performance of the ASRS. Thus, work with matched samples will be essential to providing continued information on the ASRS.

Future Directions

Replication of the current study in different contexts and with different populations will allow a greater understanding of the true nature of the diagnostic utility of the ASRS. The current study examined the ASRS within the context of children referred for potential special education services using two groups of children with developmental disabilities. Future studies should examine the ASRS in groups of children across different contexts such as typically developing children and psychiatric samples. Examination into how the ASRS performs in differentiating groups of children with an ASD from typically developing children in doctors' offices as well as psychiatric facilities and other institutions where children are routinely evaluated for various disabilities and medical diagnoses will be helpful in examining the overall potential utility of the ASRS.

Future studies should explore the older age version of the ASRS, 6-to 18-year-old, to determine how it may function within a special education setting. The types of developmental issues as well as the behavioral presentations that make it difficult to differentiate children with a potential ASD from other groups of children with a general DD may not be as prevalent in the older age group as it is with a preschool population. The utility of the 6-to 18-year-old version may be enhanced by the fact that there is not as much overlap in overt behavioral presentation between ASD and other developmental disabilities at older

ages. However, it should be noted that in the older age group the benefits of early identification and intervention will be less salient than when the child has been identified in his or her younger years.

Exploration into how the ASRS may perform with different cultural and ethnic groups will also be important to future research. This study, as well as the standardization sample, utilized a predominantly Caucasian sample, which may match the U.S. Census data and may be representative of the population in general, but may not have included enough children of certain racial or ethnic groups to capture attributes that may be unique to any respective culture. There have been differences in reported prevalence rates of ASDs in other ethnicities (Palmer, Walker, Mandell, Bayles, & Miller, 2010). These differences in prevalence may warrant further exploration into the performance of the ASRS, as well as what cut scores may be most beneficial when the instrument is used with cultural and ethnic groups outside of a predominantly Caucasian sample.

Future research should also examine how the ASRS can be used in conjunction with other measures to best identify children who may be on the autism spectrum. It should be noted that the ASRS, or any measure, should not be used in isolation in the identification of an ASD and should be used in conjunction with a battery of other assessments to assist in identifying children with potential ASDs. For example, subsequent research studies should focus on what areas of assessment (e.g., cognitive, adaptive), when used in conjunction with the ASRS, may be best at identifying and differentiating children with potential ASDs from other types of DDs. Emphasis may be placed on those areas of development that appear most salient at predicting an ASD diagnosis, such as delays in adaptive or cognitive functioning.

Utilizing the best diagnostic criteria in identifying children who may be on the autism spectrum remains essential to providing the best service to populations of children with potential ASDs. Performing utility analysis on the ASRS with more recent and upcoming criterion measures may also show benefits in terms of developing a better picture of the psychometric properties of the ASRS. As newer diagnostic criteria become available through revised versions of the Diagnostic and Statistical Manual of Mental Disorders (e.g., DSM-V) as well as updates to the Autism Diagnostic Observation Schedule (e.g., ADOS-2), using the ASRS as a predictor in studies using the most recent updates in diagnostic criteria may lead to a different overall picture of the measure's performance across a number of settings.

Conclusions

The current study set out to examine the validity of the Autism Spectrum Rating Scales (ASRS) with a preschool population referred for potential special education services. The study sought to examine aspects of the instrument's utility and determine its ability to differentiate preschool children who may be on the autism spectrum from those children having another form of developmental disability such as a speech and language impairment. The study explored how the instrument may perform with changes in cut scores and base rates, as well as in different applications and settings. Through the examination of the utility of the ASRS, it is apparent that cut scores, base rates of autism and an examination of the most accepted forms of error by context are important caveats to consider when utilizing any measure that attempts to identify children who may have an ASD.

The current study addressed the fact that when evaluating for the possibility of an ASD, it should be noted by the clinician that there is a complex interplay of a constellation of

factors that complicate the identification of children with a possible ASD, and it requires the right knowledge and assessment tools to successfully navigate this process. Increasingly, there are cost and time restrictions that minimize the optimal types of resources that can be employed in the identification of ASDs in the schools. As a result, school systems are unable to train appropriate personnel in the process of identifying ASDs and shortcuts in the assessment process are often employed. Behavior checklists increasingly serve as a means to fill in a gap and compensate for more appropriate resources, often becoming the sole measures in the identification process of children on the autism spectrum. Unfortunately, the current assessment tools that are available to assess for ASDs often have inadequate psychometric properties. As a result, misidentification may occur and possible opportunities for early intervention services may be lost by children and families.

The ASRS is a newly published behavior checklist that can be used with a preschool population of children with suspected ASDs. There are considerations that the clinician needs to make when using the ASRS. The current study has suggested that when attempting to differentiate children with an ASD from those with another type of DD, a lower cut score may be more beneficial to the clinician. Although, in general, a lower cut score may result in more false positives. In a situation where early identification and intervention will be vital for long-term benefits and positive effects, more false positives may be the more tolerable of an error to make.

The study highlighted other factors clinicians may wish to attend to when using the ASRS in attempting to identify young children on the autism spectrum in the school setting. In addition to utilizing lower cut scores in general, optimal cut scores for Parent, Teacher and Short Versions of the ASRS have all been identified through the current study. In addition to

using a lower cut score of the ASRS to successfully identify children with a potential ASD, a clinician may also wish to examine those scales that most differentiated between children on the autism spectrum and children with a general DD as identified in the current study. Furthermore, performing an item analysis with particular attention to those items identified in the current study that best differentiated children on the autism spectrum from children with a general DD can also serve to supplement a clinician in his or her identification and decision-making process. In cases of time restraint, the Short Form ASRS had comparable utility to the full Parent and Teacher Forms. Thus, a clinician could utilize the Short Form ASRS without fearing a compromise to clinical utility. In general, using the ASRS as a diagnostic tool in combination with best practice procedures as opposed to a screener appears to be the best use for the instrument.

A final consideration with regard to the ASRS is that it appears to perform differently in higher and lower base rate conditions. The ASRS performs better in populations with higher base rates of children with an ASD. This follows a general line of reasoning indicating that diagnostic measures typically discriminate best when a condition they are attempting to identify does not have a low base rate. The ability of the ASRS to perform better in higher base rate conditions indicates that the ASRS may function better in settings where there is a greater population of children with ASDs such as in special education or similar settings, as opposed to contexts with lower base rates such as pediatricians' offices. In all, despite the errors found in the current study, it appears that the ASRS has similar if not better psychometric properties when compared to other measures being used to identify ASDs.

In the increasingly complicated world of identifying children with ASDs, a clinician has many tools at his or her disposal. A clinician's approach must take into consideration the

complex interplay of factors that goes into the identification process of any particular condition. Among these factors are the utility of the diagnostic tool that may be in use, base rates of the condition being identified, tolerable types of error, and consequences and outcomes of decisions. Each of these issues has been highlighted in the current research study and placed in a context that can be applied directly to use of the ASRS in the possible identification of preschool children who may have an ASD or general DD referred for special education services.

APPENDIX A

DSM-IV-R CRITERIA FOR AUTISM SPECTRUM DISORDERS

Diagnostic Criteria for 299.00 Autistic Disorder

A. Six or more items from (1), (2), and (3), with at least two from (1), and one each from (2) and (3):

1. qualitative impairment in social interaction as manifested by at least two of the following:

    a. marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction

    b. failure to develop peer relationships appropriate to developmental level

    c. a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest)

    d. lack of social or emotional reciprocity

2. qualitative impairments in communication as manifested by at least one of the following:

    a. delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)

    b. in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others

    c. stereotyped and repetitive use of language or idiosyncratic language

    d. lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level

3. restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:

    a. encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus

    b. apparently inflexible adherence to specific, nonfunctional routines or rituals

    c. stereotyped and repetitive motor manners (e.g., hand or finger flapping or twisting, or complex whole-body movements)

    d. persistent preoccupation with parts of objects

B. Delays or abnormal functioning in at least one of the following areas with onset prior to age 3 years: (1) social interaction, (2) language as used in social communication, or (3) symbolic or imaginative play.

C. The disturbance is not better accounted for by Rettâ€™s Disorder or Childhood Disintegrative Disorder.

Diagnostic Criteria for 299.80 Asperger's Disorder

A. Qualitative impairment in social interaction as manifested by at least two of the following:

    1. marked impairment in the use of multiple nonverbal behaviors such as eye-to eye gaze, facial expression, body postures, and gestures to regulate social interaction

    2. failure to develop peer relationships appropriate to developmental level

    3. a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest to other people)

    4. lack of social or emotional reciprocity

B. Restricted repetitive and stereotyped patterns of behavior, interests and activities, as manifested by at least one of the following:

    1. encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity of focus

    2. apparently inflexible adherence to specific, nonfunctional routines or rituals

3. stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movements)

4. persistent preoccupation with parts of objects

C. The disturbance causes clinically significant impairment in social, occupational, or other important areas of functioning.

D. There is no clinically significant general delay in language (e.g., single words used by age 2 years, communicative phrases used by age 3 years).

E. There is no clinically significant delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behavior (other than in social interaction), and curiosity about the environment in childhood.

F. Criteria are not met for another specific Pervasive Developmental Disorder or Schizophrenia.

Diagnostic Criteria for 299.80 Pervasive Developmental Disorder Not Otherwise Specified

This category should be used when there is a severe and pervasive impairment in the development of reciprocal social interaction associated with impairment in either verbal or nonverbal communication skills or with the presence of stereotyped behavior, interests, and activities, but the criteria are not met for a specific Pervasive Developmental Disorder, Schizophrenia, Schizotypal Personality Disorder, or Avoidant Personality Disorder. For example, this category includes "atypical autism" – presentations that do not meet the criteria for Autistic Disorder because of late age at onset, atypical symptomatology, or subthreshold symptomatology, or all of these.

APPENDIX B

IDEA REGULATIONS PART 300(A)(300.8)

(a) General.

(1) Child with a disability means a child evaluated in accordance with Sec. 300.304 through 300.311 as having mental retardation, a hearing impairment (including deafness), a speech or language impairment, a visual impairment (including blindness), a serious emotional disturbance (referred to in this part as "emotional disturbance"), an orthopedic impairment, autism, traumatic brain injury, an other health impairment, a specific learning disability, deaf-blindness, or multiple disabilities, and who, by reason thereof, needs special education and related services.

(2)

(i) Subject to paragraph (a)(2)(ii) of this section, if it is determined, through an appropriate evaluation under Sec. 300.304 through 300.311, that a child has one of the disabilities identified in paragraph (a)(1) of this section, but only needs a related service and not special education, the child is not a child with a disability under this part.

(ii) If, consistent with Sec. 300.39(a)(2), the related service required by the child is considered special education rather than a related service under State standards, the child would be determined to be a child with a disability under paragraph (a)(1) of this section.

(b) Children aged 3 through 9 experiencing developmental delays. Child with a disability for children aged 3 through 9 (or any subset of that age range, including ages 3 through 5), may, subject to the conditions described in Sec. 300.111(b), include a child--

(1) Who is experiencing developmental delays as defined by the State and as measured by appropriate diagnostic instruments and procedures, in one or more of the following areas: physical development, cognitive development, communication development, social or emotional development, or adaptive development; and

(2) Who, by reason thereof, needs special education and related services.

(c) Definitions of disability terms. The terms used in this definition of a child with a disability are defined as follows:

(1)

(i) Autism means a developmental disability significantly affecting verbal and nonverbal communication and social interaction, generally evident before age 3, that adversely affects a child's educational performance. Other characteristics often associated with autism are engagement in repetitive activities and stereotyped movements, resistance to environmental change or change in daily routines, and unusual responses to sensory experiences.

(ii) Autism does not apply if a child's educational performance is adversely affected primarily because the child has an emotional disturbance, as defined in paragraph (c)(4) of this section.

(iii) A child who manifests the characteristics of autism after age 3 could be identified as having autism if the criteria in paragraph (c)(1)(i) of this section are satisfied.

(2) Deaf-blindness means concomitant hearing and visual impairments, the combination of which causes such severe communication and other developmental and educational needs that they cannot be accommodated in special education programs solely for children with deafness or children with blindness.

(3) Deafness means a hearing impairment that is so severe that the child is impaired in processing linguistic information through hearing, with or without amplification, that adversely affects a child's educational performance.

(4)

(i) Emotional disturbance means a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child's educational performance:

(a) An inability to learn that cannot be explained by intellectual, sensory, or health factors.

(b) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers.

(c) Inappropriate types of behavior or feelings under normal circumstances.

(d) A general pervasive mood of unhappiness or depression.

(e) A tendency to develop physical symptoms or fears associated with personal or school problems.

(ii) Emotional disturbance includes schizophrenia. The term does not apply to children who are socially maladjusted, unless it is determined that they have an emotional disturbance under paragraph (c)(4)(i) of this section.

(5) Hearing impairment means an impairment in hearing, whether permanent or fluctuating, that adversely affects a child's educational performance but that is not included under the definition of deafness in this section.

(6) Mental retardation means significantly subaverage general intellectual functioning, existing concurrently with deficits in adaptive behavior and manifested during the developmental period, that adversely affects a child's educational performance.

(7) Multiple disabilities means concomitant impairments (such as mental retardation-blindness or mental retardation-orthopedic impairment), the combination of which causes such severe educational needs that they cannot be accommodated in special education programs solely for one of the impairments. Multiple disabilities does not include deaf-blindness.

(8) Orthopedic impairment means a severe orthopedic impairment that adversely affects a child's educational performance. The term includes impairments caused by a congenital anomaly, impairments caused by disease (e.g., poliomyelitis, bone tuberculosis), and impairments from other causes (e.g., cerebral palsy, amputations, and fractures or burns that cause contractures).

(9) Other health impairment means having limited strength, vitality, or alertness, including a heightened alertness to environmental stimuli, that results in limited alertness with respect to the educational environment, that –

(i) Is due to chronic or acute health problems such as asthma, attention deficit disorder or attention deficit hyperactivity disorder, diabetes, epilepsy, a heart condition, hemophilia, lead poisoning, leukemia, nephritis, rheumatic fever, sickle cell anemia, and Tourette syndrome; and

(ii) Adversely affects a child's educational performance.

(10) Specific learning disability. (i) General. Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia.

(ii) Disorders not included. Specific learning disability does not include learning problems that are primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

(11) Speech or language impairment means a communication disorder, such as stuttering, impaired articulation, a language impairment, or a voice impairment, that adversely affects a child's educational performance.

(12) Traumatic brain injury means an acquired injury to the brain caused by an external physical force, resulting in total or partial functional disability or psychosocial impairment, or both, that adversely affects a child's educational performance. Traumatic brain injury applies to open or closed head injuries resulting in impairments in one or more areas, such as cognition; language; memory; attention; reasoning; abstract thinking; judgment; problem-solving; sensory, perceptual, and motor abilities; psychosocial behavior; physical functions; information processing; and speech. Traumatic brain injury does not apply to brain injuries that are congenital or degenerative, or to brain injuries induced by birth trauma.

(13) Visual impairment including blindness means an impairment in vision that, even with correction, adversely affects a child's educational performance. The term includes both partial sight and blindness.

(Authority: 20 U.S.C. 1401(3); 1401(30))

APPENDIX C

IDEA CLASSIFICATION CRITERIA

I. AUTISM

ELIGIBILITY CRITERIA

(1) The autism must adversely affect the student's educational performance.

(2) The student with autism must require special education and related services.

(3) The team must determine that autism is the student's primary disability, although the student may exhibit characteristics of other disability conditions such as an emotional disturbance or intellectual disability. Autism may include other conditions included in the autism spectrum, such as high functioning autism, Asperger syndrome, and pervasive developmental disorder not otherwise specified.

(4) To be eligible under this category, the student must exhibit significant impairments in verbal and/or nonverbal communication and social interaction. The student may also exhibit engagement in repetitive activities and stereotyped movements, resistance to environmental change or change in daily routines, difficulty with emotional regulation, and unusual responses to sensory experiences.

(a) Significant impairment in social interaction includes, but is not limited to:

(i) Failure to use appropriate nonverbal behaviors such as eye contact, facial expression, body postures, and other social gestures.

(ii) Failure to develop peer relationships appropriate to developmental level.

(iii) A lack of spontaneous initiation to share interests, enjoyment, or achievements with other people.

(b) Significant impairment in communication includes, but is not limited to:

(i) Delay in, or lack of, spoken language with no attempt to communicate through alternate modes such as gesture or mime.

(ii) In individuals with adequate speech: (A) An inability to initiate or sustain a conversation with others. (B) An inability to use conventions of social communication or pragmatics.

(iii) Stereotyped and repetitive use of language or peculiar language.

(iv) Lack of varied, spontaneous make-believe play, or social imitative play, appropriate to development level.

(c) Significant restricted, repetitive, and stereotyped patterns of behavior, interests, and activities includes, but is not limited to:

(i) Restricted patterns that are atypical either in intensity or focus.

(ii) Rigid adherence to specific, nonfunctional routines or rituals.

(iii) Stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movement).

(iv) Persistent preoccupation with people, events, or objects.

(d) Unusual resistance to environmental change or change in daily routines includes, but is not limited to, resistance to:

(i) New adults or students in the classroom setting, such as substitute teachers.

(ii) Changes in the arrangement of furniture.

(iii) Changes in the daily schedule of activities.

(e) Unusual responses to sensory experiences include, but are not limited to, unusual or extreme responses to:

(i) Sudden loud noises or high-pitched sounds.

(ii) Rough or highly textured surfaces or clothes touching the skin.

(iii) Bright light or significant intermittent changes in lighting.

(iv) Strong or unfamiliar tastes or smells.

EVALUATION

(1) Multiple measures (formal and informal), including an autism checklist/rating scale, must be used to assess intellectual, academic, communicative, social, and adaptive functioning.

(2) The student's prior medical and developmental history from a qualified health professional must be on record regarding specific syndromes, health concerns, medication, and any information deemed necessary for planning the student's education program.

II. DEVELOPMENTAL DELAY.

ELIGIBILITY CRITERIA

A team of qualified professionals and the student's parents determine eligibility as defined above.

(1) The team must determine that the student's primary disability is developmental delay, and not one of the other disability categories. The team should also consider whether adequate evaluation data are available to show that the student meets one of the other specific disability categories. When adequate evaluation data are available, the student must be classified in one of the other specific disability categories.

(2) The developmental delay must adversely affect the student's educational performance.

(3) The student with a developmental delay must require special education and related services.

(4) Students who are eligible for services include students who have been determined to have a significant delay or deficit in one or more of the following areas:

(a) Cognitive development.

(b) Physical/motor development.

(c) Language/speech development.

(d) Social/emotional development.

(e) Self-help skills/adaptive behavior.

(5) Significant delays are defined as:

(a) l.5 standard deviations below the mean, or at or below the 7th percentile in three areas of development.

(b) 2.0 standard deviations below the mean, or at or below the 2nd percentile in two areas of development.

(c) 2.5 standard deviations below the mean, or at or below the 1st percentile in one area of development.

EVALUATION

Multiple measures (formal and informal) must be used to assess the area(s) of suspected delay.

(1) Assessments selected must be appropriate for students ages 3 through 7 and based upon a student's sensory, motor, and communication limits.

MULTIPLE DISABILITIES

ELIGIBILITY CRITERIA

A team of qualified professionals and the student's parents determine eligibility as defined above. The team must identify the disabilities and ensure that the student meets the criteria for each of the multiple disabilities. Intellectual disabilities need not be one of the multiple disabilities identified.

(1) The multiple disabilities must adversely affect the student's educational performance.

(2) The student with multiple disabilities must require special education and related services.

EVALUATION

Multiple measures (formal and informal) must be used to assess all areas of concern. Areas to be considered include cognitive ability, academic skills, adaptive skills, language and communication, social functioning (such as self-help and independent living skills), vocational skills, and sensory/motor skills. The evaluation process is determined by the evaluation team and must include a combination of tests, interviews with those familiar with the student, and observations conducted in settings familiar to the student.

(1) Cognitive ability must be assessed by a qualified examiner. Traditional approaches to assessing cognitive ability may be of limited value for some students who are suspected of having multiple disabilities.

(2) The use of assisted and augmentative communication and motor systems must be considered during the evaluation and documented.

(3) The student's prior medical history, from a qualified health professional, must be on record if specific syndromes, special health problems (e.g., tracheotomy), medication, and long-term medical prognosis are a concern for the individual.

(4) The following sensory/motor areas must be considered for evaluation:

(a) Abnormal tactile or joint sensation,

(b) Abnormal muscle tone and movement,

(c) Lack of integration of primitive reflexes,

(d) Lack of balance or coordination,

(e) Organization of sequential motor movement,

(f) Motor skills, or

(g) A combination of any of the above.

(5) Where deficits in adaptive behavior are suspected, they must be measured and documented on standardized and/or curriculum-based assessments with input from parents and school staff.

(6) Vision and hearing must be assessed.

OTHER HEALTH IMPAIRMENT

ELIGIBILITY CRITERIA

A team of qualified professionals and the student's parents determines eligibility as defined above.

(1) The health impairment must adversely affect the student's educational performance.

(2) The student with the health impairment must require special education and related services.

(3) The team must determine that the other health impairment is the student's primary disability.

EVALUATION

(1) The student's prior medical history, from a qualified health or mental health professional, must be on record regarding specific syndromes, health concerns, medication, and any information deemed necessary for planning the student's educational program.

(2) Multiple measures (formal and informal) must be used to assess all areas of suspected deficits (e.g., educational, adaptive, behavioral, physical).

SPEECH/LANGUAGE IMPAIRMENT

ELIGIBILITY CRITERIA

A team of qualified professionals and the student's parents, including a qualified speech/language pathologist (SLP), determines eligibility as defined above.

(1) The speech or language impairment must adversely affect the student's educational performance.

(2) The student with the speech or language impairment must require special education and related services.

(3) Students who qualify in disability categories other than that of speech or language impairment may qualify for speech or language impairment services; however, in order for the student to be classified as having a speech or language impairment, the team must determine that the speech or language impairment is the student's primary disability.

(4) In order for a student whose primary home language is other than English to be eligible for classification with a speech or language impairment, the team (including an SLP) must

determine that the speech or language impairment exists in the student's primary language and is not the result of learning English as a second language.

(5) The student with an Orofacial Myofunctional Disorder or OMD (formerly called Tongue Thrust) may be served only if there is an associated speech or language impairment.

(6) Some students with mild hearing impairments may be classified as having a speech or language impairment, if the manifestation of the disability is only as a speech or language impairment and the services of a teacher of the hearing impaired are not required.

EVALUATION

Multiple measures (formal and informal) are required for a student suspected of having a speech or language impairment (primary disability or requiring related services).

(1) The student must be evaluated by a qualified SLP using assessment instruments and procedures that are appropriate for the determination and appraisal of a speech or language impairment.

(2) Documentation must be provided that indicates that the student has an impairment in listening, reasoning, and/or speaking to such a degree that special education is needed.

(3) For the student suspected of having a speech impairment, the team should consider the potential relationship of such an impairment to phonological processing and phonemic awareness.

(4) A complete battery of assessments (e.g., intellectual, physical, or adaptive behavior) may not be needed to determine that a speech or language impairment exists.

APPENDIX D

INITIAL EMAIL TO PARENTS

Dear [Parent]:

As a part of the Davis School District's preschool program, [Child's name] is eligible to participate in a new study that examines the identification of children with developmental disabilities. We will be calling you soon to determine your interest in participating. If you would like to enroll immediately or have questions or concerns, please contact the primary investigator, Sean Cunningham at 801-402-1996 during the hours of 8am and 4pm. Any voicemail is confidential and will be checked regularly in order to address your call promptly. If you do not wish to participate, please reply to this email stating such and there will be no further contact on the part of the study's personnel. Thank you for your time and interest in Davis School District's continuing efforts to enhance the educational experience of all students. Frequently asked questions are below.

What is the Focus of the Research Project?

Sometimes it may be difficult to determine what school services will best serve the needs of preschoolers who show evidence of a possible development delay. Sometimes it may also be difficult to know what type of developmental delay these children are experiencing because children with different developmental delays may look the same behaviorally. Some developmental delays with similar behavioral presentation include communication delay, intellectual delay and autism spectrum disorders. Because each of

these developmental delays may present similarly, children may be misidentified. Misidentifying children may mean the child receives educational services that may not be appropriate. Having a comprehensive assessment done during preschool has been found to be important in helping to determine whether a child has a developmental delay and what type of delay that may be. Educational services are based on the type of delay the child has including what type and how extensive services are. Davis School District is always looking for ways to improve the assessment of preschool children in order to correctly identify potential delays for the most appropriate educational supports and services. The Davis School District's Research Project on Autism Spectrum Disorders will offer a more extensive evaluation of preschoolers who have been identified as having a potential delay and possible need for additional school services.

Who will be conducting the assessments?

The assessments will be conducted by a licensed school psychologist and doctoral student at the University of Utah, Sean Cunningham. Licensed psychologist and University of Utah Professor, Dr. Elaine Clark, will provide supervision.

Is the Research Assessment Process Invasive?

No. The research project involves parents filling out some behavioral questionnaires and children participating in some tests. The questionnaires and tests examine the following areas: cognitive, language, adaptive functioning, and autism related behaviors. Parents will also be asked to participate in an interview regarding the child's developmental history and current concerns.

What will happen with the information that I provide?

All information collected during the assessments will be confidential with names of participants (including parents') separated from data. Only unidentified data will be shared with the schools and community to help improve the process of identification of preschoolers with potential developmental delays for educational services.

Who can participate?

Children ages 2 to 5 years old identified as having a need for possible additional educational services in the Davis School District.

How do I get involved in the Research Project?

Further details about the project can be obtained by contacting Sean Cunningham at the Davis School District.

Sean Cunningham, Davis School District
Email: scunningham@dsdmail.net

Washington Elementary
340 West 650 South
Bountiful, Utah 84010
(801) 402-1950

Early Learning Center (F)
115 South 200 East
Farmington, Utah 84025
(801) 402-5409

APPENDIX E

CONSENT

BACKGROUND

You and your child are being asked to take part in a research study. Before you decide, it is important for you and your child to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether you and your child would like to take part in this study. A brief summary of the purpose of the research study is explained below.

Sometimes it may be difficult to determine what school services will best serve the needs of preschoolers who show evidence of a possible developmental delay because children with different types of developmental delays may demonstrate similar behaviors. Some developmental delays with similar behavioral presentation include communication delay, intellectual delay and autism spectrum disorders. Because each of these developmental delays may present similarly, children may be misidentified. Misidentifying children may mean the child receives educational services that may not be appropriate. Having a comprehensive assessment completed when a child is young, such as during preschool, has been found to be important in helping to determine whether a child has a developmental delay and what specific type of delay that may be. Determining what educational services are appropriate and how extensive those services need to be is based on

correctly identifying the specific type of delay the child has. The Davis School District is working to improve the assessment of preschool children in order to identify potential delays correctly for appropriate educational supports and services. The Primary Investigator, Sean Cunningham, is a licensed School Psychologist in the Davis School District who will provide a more extensive evaluation of preschoolers who have been referred for an assessment due to a potential developmental delay and help determine the need for additional school services. Part of this assessment will include the Autism Spectrum Rating Scale, which is being studied in terms of its accuracy in identifying a specific type of developmental delay.

STUDY PROCEDURE

The current research project involves parents filling out behavioral questionnaires and children participating in some developmental and psychological tests. These tests are part of the typical protocol given to preschoolers who are being considered for possible special education services due to a potential developmental delay. The Autism Spectrum Rating Scale will be utilized in the study as an experimental measure that is being explored for its ability to determine if a child may have an Autism Spectrum Disorder, or other developmental delay. The total time expected to complete all the questionnaires and psychological testing will be approximately three to four hours. This time will be broken up into three smaller sessions of about one to one and a half hours. The questionnaires and tests of the current study examine the following areas: cognitive, language, adaptive functioning, and autism related behaviors.

During the first assessment session, children will be asked to participate in the Autism Diagnostic Observation Schedule (ADOS). ADOS testing takes approximately one half to

one hour to complete. The ADOS is considered the "Gold Standard" for assessing Autism and Pervasive Developmental Disorders. The ADOS is a semi-structured assessment that takes the child through a number of play tasks to evaluate communication and social interaction. Examples of these tasks include a birthday party and bubble play. Developmentally appropriate toys such as blocks, a jack-in-the-box, and miniatures (e.g., toy cars and airplanes) are also used in the tasks of the ADOS. During the second session, parents will be asked to complete two behavior checklists, the Autism Spectrum Rating Scale and the Scales of Independent Behavior – Revised. Filling out these measures takes approximately a half hour. The Autism Spectrum Rating Scale (ASRS) is a norm-referenced behavioral measure designed to identify symptoms, behaviors, and associated features of Autism Spectrum Disorders. The Scales of Independent Behavior – Revised (SIB-R) is also a norm-referenced behavioral measure that assesses adaptive behavior across several domains within the contexts of the home and community. During the third session, children will complete cognitive and language testing. This testing takes approximately one to one and a half hours to complete. The assessment used to measure cognitive functioning depends on if the child is verbal or nonverbal. Children who are verbal will participate in taking the Stanford Binet – Fifth Edition (SB5), which is an individually administered intelligence test (i.e., examiner and child work one-on-one). The SB5 assesses five domains: Fluid Reasoning (verbal and nonverbal problem solving using reasoning), Knowledge (fund of general information), Quantitative Reasoning (working and solving problems with numbers), Visual-Spatial Processing (analyzing patterns, relationships, and spatial orientation), and Working Memory (ability to store, sort and transform information in short-term memory). If the child is nonverbal, he or she will participate in the Leiter International Performance Scale

– Revised (Leiter-R), which is a nonverbal individually administered assessment of nonverbal intelligence (i.e., examiner and child work one-on-one). The Leiter-R asks the child to solve a series of tasks of which the answers do not require verbalizations. Instead, examinees utilize a series of manipulatives and cards. Language testing for the children will also be done during the third session and will consist of The Peabody Picture Vocabulary Test – Fourth Edition (PPVT-4) and the Expressive Vocabulary Test – Second Edition (EVT-2). The PPVT-4 is a standardized measure of receptive vocabulary. During the assessment, children are asked to identify a picture that represents a word given by the administrator by selecting one picture out of a series of four. The EVT-2 is a standardized measure of expressive vocabulary and word retrieval. Children are presented with a stimulus picture and asked to answer a question by the examiner, verbally give the most appropriate label, or provide a synonym for the stimulus.

All assessments will be carried out by the by the primary investigator, a licensed school psychologist, along with another similarly-trained school psychologist, intern, or licensed professional.

RISKS

The risks of this study are minimal. You may feel upset thinking about or talking about personal information related to your child's social and educational performance. These risks are similar to those you experience when discussing personal information with others. If you feel upset from this experience, you can tell the Principal Investigator, and he will tell you about resources available to help.

BENEFITS

There may be no direct benefits to you or your child for taking part in this study. Possible benefits include a more comprehensive assessment of your child's educational needs with the addition of the Autism Spectrum Rating Scale. We hope the information we get from this study may help develop a greater understanding of the assessment of children with developmental disabilities in the future.

CONFIDENTIALITY

All information collected during the assessments will be kept confidential with names of participants (including parents') separated from data and each will be stored in locked cabinets. Electronic storage will be on password-protected computers. Information collected for the study will only be available to the school's special education team and the investigators of the study. Deidentified data will be shared with the school personnel to help improve the process of identification of preschoolers with potential developmental delays for educational services. In any publications, data will be presented in a group format, to avoid identifying individuals who participate in the study.

The only exception to maintaining you and your child's confidentiality is in the event that information is disclosed that requires mandatory reporting. If you or your child discloses actual or suspected abuse, neglect, or exploitation of a child, or disabled or elderly adult, the researcher or any member of the study staff must, and will, report this to Child Protective Services (CPS), Adult Protective Services (APS), or the nearest law enforcement agency.

PERSON TO CONTACT

Primary Investigator: Questions, complaints or concerns about this study can be directed to Sean Cunningham during the hours of 8am to 4pm at 801-402-1996. If a call is placed after hours, voicemail will be checked on a continuous basis. Moreover, the voicemail is confidential (i.e., only the primary investigator has access). The faculty supervisor for the research project is Elaine Clark, Ph.D., and can be reached during the hours of 9am to 5pm at 801-581-7148. Voicemail at the number is checked regularly. Either of the previously mentioned individuals may be contacted in the case you or your child feel harmed by the research.

Institutional Review Board: Contact the Institutional Review Board (IRB) if you have questions regarding your child's rights as a research participant. Also, contact the IRB if you have questions, complaints or concerns which you do not feel you can discuss with the investigator. The University of Utah IRB may be reached by phone at 801 581-3655 or by e-mail at irb@hsc.utah.edu.

Research Participant Advocate: You may also contact the Research Participant Advocate (RPA) by phone at 801 581-3803 or by email at participant.advocate@hsc.utah.edu.

VOLUNTARY PARTICIPATION

It is up to you to decide whether you and your child take part in this study. Refusal to participate or the decision to withdraw from this research will involve no penalty or loss of benefits to which you are otherwise entitled. This will not affect your relationship with the investigator or with Davis School District.

COSTS AND COMPENSATION TO PARTICIPANTS

There are no costs or compensation for participating in this study.

CONSENT

By signing this consent form, I confirm I have read the information in this parental permission form and have had the opportunity to ask questions. I will be given a signed copy of this parental permission form. I voluntarily agree to allow my child to take part in this study.

_____

Child's Name

_____

Parent/Guardian's Name

_____ _____

Parent/Guardian's Signature Date

_____

Relationship to Child

_____

Name of Researcher or Staff

_____ _____

Signature of Researcher or Staff Date

APPENDIX F

DESCRIPTIVE STATISTICS OF THE PARENT AND TEACHER ASRS

Scale	Parent ASRS		Teacher ASRS	
	ASD	DD	ASD	DD
	(*n* = 37)	(*n* = 30)	(*n* = 25)	(*n* = 28)
Total	69.35(8.42)	66.07(11.32)	66.28(13.14)	62.00(9.98)
Social Communication	68.59(7.73)	62.20(9.22)	69.44(12.89)	62.82(10.55)
Unusual Behaviors	64.27(8.80)	65.73(11.78)	60.32(15.15)	61.75(13.29)
DSM-IV-TR Criteria	70.11 (9.02)	66.41(12.21)	66.28(12.97)	62.57(9.61)
Peer Socialization	72.16 (8.53)	65.50(9.87)	67.04 (8.43)	62.00(9.72)
Adult Socialization	67.19(10.01)	63.47(10.44)	65.24(11.24)	62.61(9.59)
Social/Emotion Reciprocity	65.30(9.26)	61.60(12.07)	65.56(12.05)	59.00(9.14)
Atypical Language	58.46(10.48)	60.40(8.72)	62.36(16.23)	56.32(9.65)
Sterotypy	65.41(9.58)	60.90(11.04)	58.48(10.85)	57.50(12.76)
Behavioral Rigidity	64.65(11.08)	66.03(13.53)	57.24(13.20)	60.57(12.89)
Attention/Self-Regulation	62.84(9.24)	59.87(11.76)	60.80(9.36)	57.29(11.08)

Note: ASRS = Autism Spectrum Rating Scales; ASD = Autism Spectrum Disorder; DD = Developmental Delay

APPENDIX G

DESCRIPTIVE STATISTICS OF THE DEVELOPMENTAL VARIABLES

Variable	ASD	DD
	(*n* = 35)	(*n* = 30)
IQ	75.29(26.84)	94.00(20.91)**
Adaptive	62.74(25.89)	90.33(25.69)***
Receptive Language	54.72(29.23)	84.86(29.25)***
Expressive Language	62.16(23.61)	80.76(28.86)**

Note: IQ = Stanford-Binet – Fifth Edition or Leiter-R; Adaptive = Scales of Independent Behavior – Revised; PPVT-4 = Peabody Picture Vocabulary Test – Fourth Edition; EVT-2 = Expressive Vocabulary Test – Second Edition; ASD = Autism Spectrum Disorder; DD = Developmental Delay.
* $p < .05$
** $p < .01$
*** $p < .001$

APPENDIX H

ITEM POINT BISERIAL CORRELATIONS AND MEANS

OF ASD AND DD GROUPS FOR PARENT ASRS

Item	*r*	ASD Mean	DD Mean
1	0.05	0.91	0.81
2	-0.10	1.44	1.70
3	0.16	2.25	1.89
4	0.22	1.69	1.30
5	0.19	1.84	1.41
6	-0.58***	0.59	2.07
7	0.47***	2.03	0.89
8	-0.09	2.50	2.70
9	0.00	2.56	2.56
10	-0.07	2.41	2.56
11	-0.10	2.22	2.48
12	-0.15	0.94	1.30
13	0.16	1.81	1.48
14	0.37**	2.97	2.19
15	0.26*	2.94	2.30
16	0.23	1.88	1.44
17	-0.20	1.28	1.70
18	0.40**	1.97	1.04
19	0.08	2.53	2.33
20	-0.09	2.25	2.44
21	0.28*	1.75	1.19
22	0.22	2.50	1.89
23	0.01	2.53	2.52
24	0.19	2.94	2.56
25	0.22	2.00	1.59
26	-0.43**	0.66	1.70
27	0.00	1.94	1.93
28	0.33*	2.59	1.74
29	0.47***	3.06	2.00
30	-0.01	2.53	2.56

Item	*r*	ASD Mean	DD Mean
31	-0.24	1.56	2.07
32	-0.01	2.16	2.19
33	0.05	1.75	1.63
34	0.00	2.75	2.74
35	0.03	1.31	1.26
36	-0.03	1.13	1.19
37	-0.02	1.44	1.48
38	0.08	2.56	2.41
39	0.36**	3.06	2.22
40	0.35**	1.97	1.30
41	-0.44**	0.72	1.85
42	0.12	2.35	2.00
43	0.20	2.41	2.00
44	0.30**	2.65	1.89
45	0.21	1.59	1.11
46	0.09	2.09	1.85
47	-0.04	2.00	2.11
48	-0.10	1.88	2.15
49	0.29*	2.34	1.70
50	0.24	2.63	2.15
51	0.28*	2.69	2.11
52	0.14	2.53	2.19
53	0.00	1.68	1.67
54	0.07	1.47	1.33
55	0.16	1.81	1.48
56	-0.01	2.50	2.52
57	0.22	1.81	1.37
58	-0.25	2.16	2.74
59	-0.16	1.10	1.52
60	0.06	2.00	1.85
61	0.23	2.41	2.04
62	0.19	2.65	2.23
63	0.16	2.92	2.62
64	0.00	0.92	0.92
65	0.08	1.08	0.88
66	0.34*	1.04	0.31

Item	*r*	ASD Mean	DD Mean
67	0.40**	2.35	1.19
68	0.21	0.79	0.35
69	0.00	1.00	1.00
70	0.15	1.79	1.38

Note: ASRS = Autism Spectrum Rating Scales; ASD = Autism Spectrum Disorder; DD = Developmental Delay
* $p < .05$
** $p < .01$
*** $p < .001$

APPENDIX I

ITEM POINT BISERIAL CORRELATIONS AND MEANS OF ASD AND DD GROUPS FOR TEACHER ASRS

Item	*r*	ASD Means	DD Means
1	-0.09	1.25	1.42
2	-0.10	0.75	1.00
3	0.41**	2.92	2.12
4	0.33*	2.42	1.58
5	0.25	2.00	1.50
6	-0.44**	0.79	1.77
7	0.50***	2.04	0.88
8	-0.29*	1.38	2.12
9	-0.14	1.63	2.00
10	-0.07	1.67	1.85
11	-0.28*	0.88	1.54
12	-0.18	0.42	0.73
13	0.17	2.00	1.69
14	0.43**	3.13	2.31
15	0.38**	3.13	2.19
16	0.19	2.17	1.73
17	0.12	1.88	1.58
18	0.31*	2.75	2.00
19	0.29*	3.00	2.46
20	-0.14	1.50	1.88
21	0.15	1.63	1.31
22	0.34*	2.63	1.62
23	-0.04	2.46	2.54
24	0.22	2.75	2.19
25	0.33*	2.08	1.38
26	-0.35	0.50	1.35

Item	*r*	ASD Means	DD Means
28	0.38**	2.92	2.04
29	0.55***	3.25	1.96
30	0.01	2.29	2.27
31	0.08	1.96	1.77
32	0.22	2.54	2.00
33	0.09	1.79	1.58
34	0.11	2.54	2.27
35	0.22	1.75	1.31
36	0.19	1.67	1.31
37	0.21	1.42	0.96
38	0.16	2.67	2.35
39	0.02	1.63	1.58
40	0.32*	2.38	1.73
41	-0.14	0.79	1.12
42	0.14	1.88	1.42
43	0.33*	2.46	1.73
44	0.31*	2.54	1.77
45	-0.05	1.29	1.42
46	-0.09	1.33	1.58
47	-0.18	1.13	1.58
48	0.03	0.96	0.88
49	0.16	2.38	1.96
50	0.29*	3.04	2.46
51	0.29*	2.79	2.08
52	0.38**	3.00	2.12
53	0.38**	1.67	0.69
54	0.30*	2.25	1.62
55	0.16	1.92	1.62
56	-0.17	1.25	1.65
57	0.21	1.74	1.38
58	-0.22	1.63	2.15
59	-0.01	0.75	0.77

Item	*r*	ASD Means	DD Means
61	0.27	2.71	2.12
62	0.19	2.65	2.23
63	0.16	2.92	2.62
64	0.00	0.92	0.92
65	0.08	1.08	0.88
66	0.34*	1.04	0.31
67	0.40**	2.35	1.19
68	0.21	0.79	0.35
69	0.00	1.00	1.00
70	0.15	1.79	1.38

Note: ASRS = Autism Spectrum Rating Scales; ASD = Autism Spectrum Disorder; DD = Developmental Delay
* $p < .05$
** $p < .01$
*** $p < .001$

REFERENCES

Allen, C.W., Silove, N., Williams, K., & Hutchins, P. (2007). Validity of the social communication questionnaire in assessing risk of autism in preschool children with developmental problems. *Journal of Autism and Developmental Disorders, 37,* 1272-1278.

Allen, R. A., Robins, D. L., & Decker, S. L. (2008). Autism spectrum disorders: Neurobiology and current assessment practices. *Psychology in the Schools, 45,* 905-917.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (Revised 3rd ed.). Washington, DC: Author.

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

The Autism Genome Project Consortium. (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature Genetics, 39,* 319–328.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study, *Psychological Medicine, 25,* 63-77.

Baron-Cohen, S., Cox, A., Baird, G., Sweettenham, J., & Nightingale, N. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry, 168*(2), 158–163.

Ben Itzchak, E., & Zachor, D. A. (2011). Who benefits from early intervention in autism spectrum disorders? *Research in Autism Spectrum Disorders, 5*, 345-350.

Bertone, A., Hanck, J., Kogan, C., Chaudhuri, A., & Cornish, K. (2010). Using perceptual signatures to define and dissociate condition-specific neural etiology: Autism and Fragile X Syndrome as model conditions, *Journal of Autism and Developmental Disorders, 40,* 1531-1540.

Berument, S., Rutter, M., & Lord, C., Pickles, A., Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *The British Journal of Psychiatry, 175*, 444–451.

Berument, S. K., Rutter, M., Lord, A., Pickles, A., & Bailey, B. (1999). Social Communications Questionnaire. Australia: Australian Council for Educational Research.

Bruininks R. H., Woodcock R.W., Weatherman R. F., & Hill B. K. (1996). *Scales of Independent Behavior-Revised*. Chicago, IL: Riverside.

Campbell, J. (2005). Diagnostic assessment of Asperger's disorder: A review of five third-party rating scales. *Journal of Autism and Developmental Disorders, 35*, 25–35.

Centers for Disease Control and Prevention. (2007). *Prevalence of autism spectrum disorders* Washington, DC: U.S. Government Printing Office.

Center for Disease Control and Prevention. (2012). Prevalence of autism spectrum disorders — Autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Morbidity and Mortality Weekly Report, Surveillance Summaries, 61*,1-19.

Chakrabarti, S., & Frombone, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. *American Journal of Psychiatry, 162*, 1133 - 1141.

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology, 33*, 781-789.

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1998). An experimental investigation of social-cognitive abilities in infants with autism: Clinical implications. *Infant Mental Health Journal, 19*, 260–275.

Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007) Autism spectrum disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry, 48,* 128–138.

Chawarska, K., Paul, R., Klin, A., Hannigen, S., Dichtel, L.E., & Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 62-72.

Cicchetti, D. V., Lord, C., Koenig, K., Klin, A., & Volkmar, F. R. (2008). Reliability of the ADI-R: Multiple examiners evaluate a single case. *Journal of Autism and Developmental Disorders, 38*, 764-770.

Cohen, I. L., Gomez, T. R., Gonzalez, M. G., Lennon, E. M, Karmel, B. Z., & Gardner, J. (2010). Parent PDD behavior inventory profiles of young children classified according to Autism Diagnostic Observation Schedule-Generic and Autism Diagnostic Interview-Revised criteria. *Journal of Autism and Developmental Disorders, 40*, 246-254.

Coolican, J., Bryson, S. E., & Zwaigenbaum, L. (2008). Brief report: Data on the Stanford-Binet Intelligence Scales (5th ed.) in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 38*, 190-197.

Cox, A., Klein, K., Charman, T., Baird, G., Baron Cohen, S., Swettenham, J. et al. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 719–732.

Deprey, L., & Ozonoff, S. (2009). Assessment of comorbid psychiatric conditions in autism spectrum disorders. In S. Goldstein, J. Naglieri, & S. Ozonoff (Eds.). *Assessment of autism spectrum disorders* (pp. 290-317.) New York, NY: Guilford Press.

Douglas, K. S., Otto, R. K., & Borum, R. (2003). Clinical forensic psychology. In I. B. Weiner (Ed.), *Handbook of psychology* (pp. 189 - 212). Hoboken, NJ: John Wiley & Sons, Inc.

Dover, C., & LeCouteur, A. L. (2007). How to diagnose autism. *Archives of Disease in Childhood, 92*, 540-545.

Dunn, L. M., & Dunn D.M. (2006). *Peabody Picture Vocabulary Test* (4th ed). Toronto, Ontario: Pearson Canada Assessment Inc.

Essex-Sorlie, D. (1995). *Medical biostatistics* and *epidemiology*. Norwalk, CT: Appleton & Lange.

Felder, S. & Mayrhofer, T. (2011): *Medical decision making – A health economic primer*, Berlin: Springer.

Fernell, E., Hedvall, Å., Norrelgen, F., Eriksson, M., Höglund-Carlsson, L., Barnevik-Olsson, M., et al., (2010). Developmental profiles in preschool children with autism spectrum disorders referred for intervention. *Research in Developmental Disabilities, 31*, 790-799.

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Jr., Dawson, G., et al. (2000). Practice parameter: Screening and diagnosis of autism: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology, 55*, 468-479.

Fletcher, R. W., & Fletcher, S. W. (2005). *Clinical epidemiology: The essentials.* Baltimore, MD: Lippincott Williams & Wilkins.

Fombonne E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry, 66*, 3– 8.

Ganz, M.L. (2007). The lifetime distribution of the incremental societal costs of autism. *Archives of Pediatric & Adolescent Medicine, 161,* 343-349.

Gardner, H., Spiegelman, D., & Buka, S. L. (2009). Prenatal risk factors for autism: comprehensive meta-analysis. *The British Journal of Psychiatry, 195*, 7–14.

Gilliam, J. E. (1995). *Gilliam Autism Rating Scale.* Austin, TX: Pro-Ed.

Gilliam, J. E. (2001). *Gilliam Asperger's Disorder Scale*. Austin, TX: Pro-Ed.

Gilliam, J. E. (2003). *Gilliam Asperger's Disorder Scale*. Austin, TX: Pro-Ed.

Gilliam, J. (2006). *GARS-2: Gilliam Autism Rating Scale–Second Edition*. Austin, TX: PRO-ED.

Glaros, A. G., & Kline, R. B. (1988). Understanding the accuracy of tests with cutting scores: The sensitivity, specificity, and predictive value model. *Journal of Clinical Psychology, 44,* 1013-1023.

Goldstein, S., & Naglieri, J. A. (2010). *Autism Spectrum Rating Scale*. Tonawanda¸ON: MHS.

Goldstein, S., Naglieri, J.A., & Ozonoff, S. (2009). *Assessment of autism spectrum disorders*. New York, NY: Guilford Press.

Goldstein, S., Ozonoff, S., Cook, A., & Clark, E. (2009). Alternative methods, challenging issues, and best practices in the assessment of autism spectrum disorders. In S. Goldstein, J. A. Naglieri, & S. Ozonoff (Eds.). *Assessment of autism spectrum disorders* (pp. 358-372). New York, NY: Guilford Press.

Granpeesheh, D., Dixon, D. R., Tarbox, J., Kaplan, A. M., & Wilke, A. E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3,* 1014-1022.

Gray, K., M., Tonge, B. J., & Sweeney, D. J. (2008). Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: Evaluating diagnostic validity. *Journal of Autism and Developmental Disorders, 38*, 657-667.

Gupta, V. B., Hyman, S. l., Johnson, C. P., Bryant, J., Byers, B, Kallen, R. et al., (2007). Identifying children with autism early? *Pediatrics, 119*, 152-153.

Individuals with Disabilities Education Improvement Act (2004), 20 U.S.C. §1400.

Ioannidis, J. P. A., & Tatsioni, A. (2010). Biostatistics: Meta-analysis. In J. Killewo, H. K. Heggenhougen, & S. R. Quah (Eds.), *Epidemiology and demography in public health* (pp. 440-449). San Diego, CA: Elsevier Inc.

Jacobson, J. W., Mulick, J. A., & Green, G. (1998). Cost-benefit estimates for early intensive behavioral intervention for young children with autism - General model and single state case. *Behavioral Interventions, 13*, 201-226.

Johnson, E. S., Jenkins, J. R., & Petscher, Y. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research & Practice, 24*, 174-185.

Johnson, C. P., Myers, S. M., & Council on Children with Disabilities, (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics, 120,* 1183-1215.

Klin, A., Lang, J., Cicchetti, D. V., & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders, 30,* 163-167.

Lainhart, J. E. (1999). Psychiatric problems in individuals with autism, their parents and siblings. *International Review of Psychiatry, 11*, 278-298.

Landrigan, P. J. (2011). Environment and autism. In E. Hollander, B. Kolevzon, J. T. Coyle (Eds.), *Textbook of autism spectrum disorders* (pp. 247-264). Arlington, VA: American Psychiatric Publishing, Inc.

Landry, S. H., & Loveland, K. A. (1988). Communication behaviors in autism and developmental language delay. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 29,* 621–634.

Lecavalier, L. (2005). An evaluation of the Gilliam Autism Rating Scale. *Journal of Autism and Developmental Disorders, 35,* 795-805.

Le Couteur, A., Haden, G., Hammal, D., & McConachie, H. (2008). Diagnosing autism spectrum disorders in pre-school children using two standardized assessment instruments: The ADI-R and the ADOS. *Journal of Autism and Developmental Disorders, 38,* 362-372.

Lee, L., David, A. B., Rusyniak, J., Landa, R., & Newschaffer, C. J. (2007). Performance of the Social Communication Questionnaire in children receiving preschool special education services. *Research in Autism Spectrum Disorders, 1*, 126-138.

Leonard, H., Dixon, G., & Whitehouse, A. J. O. (2010). Unpacking the complex nature of the autism epidemic. *Research in Autism Spectrum Disorders, 4*, 548-554.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36*, 1–18.

Lord, C., & Pickles, A. (1996). Language level and nonverbal social-communicative behaviors in autistic and language delayed children. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 1542–1550.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *ADOS. Autism Diagnostic Observation Schedule*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal, 14*, 234–252.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurn, A, & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63,* 694-701.

Losh, M., Sullivan, P. F., Trembath, D., & Piven, J. (2008). Current developments in the genetics of autism: From phenome to genome. *Journal of Neuropathology and Experimental Neurology, 67*, 829–837.

Matson, J. L. (2007). Current status of differential diagnosis for children with autism spectrum disorders. *Research in Developmental Disabilities, 28*, 109-118.

Matson, J. L., Nebel-Schwalm, M., & Matson, M. L. (2007). A review of methodological issues in the differential diagnosis of autism spectrum disorders in children. *Research in Autism Spectrum Disorders, 1*, 38-54.

Mazefsky, C., A., & Oswald, D. P (2006). The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting. *Autism, 10,* 533-549.

McFall, R. M. (2005). Theory and utility - Key themes in evidence-based assessment: Comment on the special section. *Psychological Assessment, 17*, 312-323.

McGee, S. (2002). Simplifying likelihood ratios. *Journal of General Internal Medicine, 17,* 647–650.

Metz, C. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, 8*, 283-298.

Mildenberger, K., Sitter, S., Noterdaeme, M., & Amorosa, H. (2001). The use of the ADI-R as a diagnostic tool in the differential diagnosis of children with infantile autism and children with a receptive language disorder. *European Child & Adolescent Psychiatry, 10,* 248-255.

National Research Council. (2001). Educating children with autism. Committee on Education and Interventions for Children with Autism. Washington, DC: Division of Behavioral and Social Sciences and Education. National Academy Press.

Newschaffer, C. J., Falb, M.D., & James G. G. (2005). National autism prevalence trends from United States special education data, *Pediatrics, 115*, 277-282.

Noterdaeme, M., Sitter, S., Mildenberger, K., & Amorosa, H. (2000). Diagnostic assessment of communicative and interactive behaviours in children with autism and receptive language disorder. *European Child & Adolescent Psychiatry, 4*, 295-300.

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*, 523–540.

Palmer, R. F., Walker, T., Mandell, D., Bayles, B., & Miller, C. S. (2010). Explaining low rates of autism among Hispanic schoolchildren in Texas. *American Journal of Public Health, 100*, 270-272.

Perry, A., Condillac, R. A., Freeman, N. L., Dunn-Geier, J., & Belair, J. (2005). Multi-site Study of the Childhood Autism Rating Scale (CARS) in five clinical groups of young children. *Journal of Autism and Developmental Disorders, 35*, 625-634.

Rice, C. (2007). Prevalence of autism spectrum disorders - Autism and developmental disabilities monitoring network, 14 sites, United States, 2002. *Mortality and Morbidity Weekly Report, 56* (SS-1) 12-28.

Rice, C. (2009). National Center on Birth Defects and Developmental Disabilities, CDC. Prevalence of Autism spectrum disorders – Autism and developmental disabilities monitoring network, United States, 2006. Published online December 18, 2009. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5810a1.htm.

Rickards, A. L., Walstab, J. E., Wright-Rossi, R. A., Simpson, J., & Reddihough, D. S. (2009). One- year follow-up of the outcome of a randomized controlled trial of a home-based intervention programme for children with autism and developmental delay and their families. *Child Care, Health & Development, 35,* 593-602.

Robins, D., Fein, D., & Barton, M. (1999). *The Modified Checklist for Autism in Toddlers (M-CHAT)*. Storrs, CT: University of Connecticut.

Rogers, S. J. (1996). Brief Report: Early intervention in autism. *Journal of Autism and Development Disorders, 26,* 243-24.

Rogers, S. J. (1998). Empirically supported comprehensive treatments for young children with autism. *Journal of Clinical Child Psychology, 27*, 168-179.

Rogers, S. J., Hepburn, S., Stackhouse, T., & Wehner, E. (2003). Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. *Journal of Autism and Developmental Disorders, 33*, 631–642.

Roid, G. (2003a). *Stanford-Binet Intelligence Scales (5th ed.)*. Itasca, IL: Riverside Publishing.

Roid, G. H. (2003b). *Stanford–Binet Intelligence Scales (5th ed.): Technical manual*. Itasca, IL: Riverside Publishing.

Roid, G. H., & Miller, L.J. (1997). *Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting.

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.

Saemundsen, E., Magnusson, P., Smari, J., & Sigurdardottir, S. (2003). Autism Diagnostic Interview-Revised and the Childhood Autism Rating Scale: Convergence and discrepancy in diagnosing autism. *Journal of Autism and Developmental Disorders, 33*, 319–328.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The Childhood Autism Rating Scale (CARS)*. Los Angeles, CA: Western Psychological Services.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *The Childhood Autism Rating Scale (CARS) - Second Edition.* Los Angeles, CA: Western Psychological Services.

Shattuck, P.T. (2002). *National trends in the prevalence of autism among children in the public schools.* Paper presented at the 35th Annual Gatlinburg Conference on Research and Theory in Intellectual and Developmental Disabilities, San Diego, CA.

Sikora, D. M., Hall, T. A., Hartley, S. L., Gerrard-Morris, A. E., & Cagle, S. (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL and GARS. *Journal of Autism and Developmental Disorders, 38*, 440-448.

Smits, N., Smit, F., Cuijpers, P. & De Graaf, R. (2007). Using decision theory to derive optimal cut-off scores of screening instruments: An illustration explicating costs and benefits of mental health screening. *International Journal of Methods in Psychiatric Research, 16*, 219–229.

Smith, T., Groen, A.D., & Wynn, J.W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal of Mental Deficiency*, *105*, 269-285.

South, M., Williams, B. J., Mcmahon, W. M., Owley, T., Filipek, P. A., Shernoff, E., et al. (2002). Utility of the Gilliam Autism Rating Scale in research and clinic populations. *Journal of Autism and Developmental Disorders, 32*, 593–599.

Spitalnic, S. (2004). Test properties 2: Likelihood ratios, Bayes' formula, and receiver operating characteristic curves. *Hospital Physician, 40*, 53-58.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests (3rd ed.)*, New York, NY: Oxford Univ. Press.

Trillingsgaard, A., Sorensen, E. U., Nemec, G., & Jorgensen, M. (2005). What distinguishes autism spectrum disorders from other developmental disorders before the age of four years? *European Journal of Child and Adolescent Psychiatry, 14*, 65–72.

Tryon, P. A., Mayes, S. D., Rhodes, R. L., & Waldo, M. (2006). Can Asperger's disorder be differentiated from autism using DSM-IV criteria? *Focus on Autism and Other Developmental Disabilities, 21,* 2-6.

Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism, 10*, 243-265.

U. S. Department of Education. (2004). *IDEA data documentation*. Retrieved July 16th, 2011, from http://www.ideadata.org/documents.asp.

U.S. Department of Education, National Center for Education Statistics (2010). *Digest of Education Statistics.*

Ventola, P., Kleinman, J., Pandey, J., Wilson, L., Esser, E., Boorstein, H., et al. (2007). Differentiating between autism spectrum disorders and other developmental disabilities in children who failed a screening instrument for ASD. *Journal of Autism and Developmental Disorders, 37,* 425-436.

Verstag, B. (2004). National autism summit charts a path through a scientific, clinical wilderness. *Journal of the American Medical Association, 29*, 29–31.

Vig, S., & Jedrysek, E. (1999). Autistic features in young children with significant cognitive impairment: Autism or mental retardation? *Journal of Autism and Developmental Disorders, 29,* 235-248.

Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J. T., Abrahams, B. S., et al. (2009). Common genetic variants on 5p14.1 associated with autism spectrum disorders. *Nature, 459,* 528- 533.

Ward, S. L., & Gilmore, L. (2010). The Autistic Behavioural Indicators Instrument (ABII): Development and instrument utility in discriminating autistic disorder from speech and language impairment and typical development. *Research in Autism Spectrum Disorders, 4*, 28-42.

Wiggins, L.D., Baio, J., & Rice, C. (2006). Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. *Journal of Developmental and Behavioral Pediatrics, 27*, S79 –S87.

Wiggins, J. S. (1973). *Personality and prediction*. Reading, MA: Addison-Wesley.

Wiggins, L. D., Bakeman, R., Adamson, L. B., & Robins, D. L. (2007). The utility of the Social Communication Questionnaire in screening for autism in children referred for early intervention. *Focus on Autism and Other Developmental Disabilities, 22*, 33-38.

Williams, K. T. (2007). *Expressive Vocabulary Test (2nd ed.).* Toronto, Ontario: Pearson Canada Assessment Inc.

Williams, M. E., Atkins, M., & Soles, T. (2009). Assessment of autism in community settings: Discrepancies in classification. *Journal of Autism and Developmental Disorders, 39*, 660–669.

Youden, W.J. (1950). Index for rating diagnostic tests. *Cancer,* 3, 32-35.