

IDENTIFYING AUTISM SPECTRUM DISORDER IN A CLINICAL SAMPLE OF
PRESCHOOL-AGED CHILDREN USING THE
BASC-3 PARENT RATING SCALES

APPROVED BY SUPERVISORY COMMITTEE

Katherine Bellone, Ph.D., BCBA

James LePage, Ph.D.

Aleksandra Foxwell, Ph.D.

Corinne Anton, Ph.D., ABPP

Stephen Elliott, M.D.

ACKNOWLEDMENTS

I would like to thank my committee chair, Dr. Katherine Bellone, and my committee supervisor, Dr. James LePage, for their ongoing mentorship and support through this long process. Thank you both for your unwavering encouragement, guidance, and willingness to step-in and help with whatever was needed. You have helped me grow both personally and professionally, while consistently allowing room for me to find my own path through this dissertation journey. I will always appreciate your kindness and support.

I would also like to thank my remaining committee members, Dr. Corinne Anton, Dr. Stephen Elliott, and Dr. Aleksandra Foxwell, for their steadfast support since this venture began. Thank you for contributing your unique sets of clinical skills and research expertise to the betterment of this project and my training. Additionally, I would like to acknowledge and thank my previous mentor, Dr. Stormi White, for helping me cultivate a passion for working with young children and assisting me in the design and creation of this research. To my clinical supervisor, Dr. Celia Heppner, thank you for always being behind the scenes of my training process offering encouragement, bolstering my confidence when needed, and guiding my professional development.

To my close family and friends, thank you for being my biggest and most important cheerleaders. To my parents, thank you for cultivating in me passion, grit, and persistence. All that I have accomplished and overcome has been fueled by your unconditional love and belief in me – what an incredible gift to give to your child; I am forever grateful. To my best friend and husband, Michael, no words of acknowledgement would ever be able to adequately express what you have given to me in support and love throughout these years, but I know that you know what is in my heart. We will always be for each other.

Lastly, to my daughter Evangeline – “beautiful girl, you were made to do hard things so believe in yourself.”

IDENTIFYING AUTISM SPECTRUM DISORDER
IN PRESCHOOL-AGED CHILDREN USING
THE BASC-3 PARENT RATING SCALES

by

SAVANNA SABLICH DULEY, B.S.

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2020

Copyright

By

Savanna Sablich Duley, 2020

All Rights Reserved

IDENTIFYING AUTISM SPECTRUM DISORDER USING THE BASC-3
PARENT RATING SCALES

Publication No. _____

Savanna Sablich Duley, B.S.

The University of Texas Southwestern Medical Center at Dallas, 2020

Supervising Professors: Katherine Bellone, PhD, BCBA; James LePage, PhD

ABSTRACT

The Behavior Assessment System for Children, Third Edition (BASC-3) is a broadband behavioral rating scale commonly used in medical and educational settings to assess a variety of emotional and behavioral difficulties in preschool age children. The DSD content scale on the BASC is intended to measure impairments in a child's social skills, communication, interests, and activities. Use of the scale has been suggested to improve early identification efforts in the areas of developmental screening and diagnosis of autism spectrum disorders (ASD). Previous research investigating the DSD scale reported evidence

for the scale's ability to aid in the identification of preschoolers with developmental delays, and discriminate between children diagnosed with ASDs, other diagnoses, and those who were typically developing; however the DSD scale has not been re-validated in clinical preschool populations using the updated BASC edition, the BASC-3. The current study examined whether T-Scores on the DSD content scale on BASC-3 Parent Rating Scales, Preschool Form could identify preschool-aged children diagnosed with an ASD and meaningfully differentiate these children from those diagnosed with other developmental delays. DSD T-Scores were generated for each participant using several different normative scoring comparison groups and compared across scoring methods to fully examine the utility of the DSD scale specific to a clinical sample. Results indicated that the DSD scale was able to effectively identify and confirm the presence of symptoms related to developmental social disorders among all participants; however, the scale was unable to distinguish between preschoolers diagnosed with ASD versus those with other developmental delays.

TABLE OF CONTENTS

CHAPTER ONE

Statement of the Problem.....	1
-------------------------------	---

CHAPTER TWO

Review of the Literature	6
Summary and Aims	26
Research Questions and Hypotheses	28

CHAPTER THREE

Methodology	31
-------------------	----

CHAPTER FOUR

Results.....	44
--------------	----

CHAPTER FIVE

Discussion.....	64
-----------------	----

LIST OF TABLES

TABLE ONE	Appendix A, 78
TABLE TWO	Appendix A, 80
TABLE THREE	46
TABLE FOUR.....	48
TABLE FIVE	52
TABLE SIX	54
TABLE SEVEN	56
TABLE EIGHT	61
TABLE NINE	62
TABLE TEN	63

LIST OF FIGURES

FIGURE ONE	60
FIGURE TWO	69
FIGURE THREE	69
FIGURE FOUR	70

LIST OF APPENDICES

APPENDIX A	78
------------------	----

LIST OF ABBREVIATIONS

ASD – Autism Spectrum Disorder

ADDM – Autism and Developmental Disabilities Monitoring

CDC – Center of Disease Control

APA – American Psychiatric Association

DSM – Diagnostic and Statistical Manual of Mental Disorders

PDD – Pervasive Developmental Disorder

RRB – Restrictive, Repetitive Behaviors

BASC – Behavior Assessment System for Children

BASC-3 PRS-P – Behavior Assessment System for Children, Third Edition, Parent

Rating Scale, Preschool

CBCL – Childhood Behavior Checklist

DSD – Developmental Social Disorders

M-CHAT – Modified Checklist for Autism in Toddlers

AAP – American Academy of Pediatrics

MSEL – Mullen Scales of Early Learning

ADOS-2 – Autism Diagnostic Observation Schedule, Second Edition

ADI-R – Autism Diagnostic Interview, Revised

Leiter-3 – Leiter Performance Scales, Third Edition

SB-5 – Stanford Binet, Fifth Edition

NVDQ – Nonverbal Developmental Quotient

NVIQ – Nonverbal Intelligence Quotient

VRae – Visual Reception Age Equivalent

FMae – Fine Motor Age Equivalent

SA score – Social Affect score (ADOS-2)

RRB score – Restrictive, Repetitive Behaviors score (ADOS-2)

CHAPTER ONE

Introduction

Statement of the Problem

Early identification of children with autism spectrum disorder (ASD) is crucial, as it often initiates access to targeted early interventions that can improve future outcomes (Koegel, Koegel, Ashbaugh, & Bradshaw, 2014; Zwaigenbaum et al., 2015a). Evidence from epidemiological studies in the U.S. document that 1 in 59 children are diagnosed with ASD and the average age of a child at initial diagnosis is around 4-years-old (Baio et al., 2018; Christensen et al., 2016; Christensen et al., 2019). Given that a culmination of diagnostic research has collectively found evidence to suggest that ASD can be reliably diagnosed in children as early as 18 months (Ozonoff et al., 2015), there remains substantial room for improvement in early identification screening and diagnostic practices in order to narrow the gap between the onset of symptoms and initial diagnosis.

Diagnosing ASD in early childhood, specifically in preschool-aged children, can be difficult due to the variety and complexity of symptom presentations, as well as the presence of other challenging behaviors that co-occur within this early developmental period (S. P. White, Weitlauf, & Warren, 2012). Emotional and behavioral problems in early childhood are often misattributed to developmental appropriateness, temperament differences, or are considered transitory in nature (Bagner, Rodriguez, Blake, Linares, & Carter, 2012). Further, behaviors exhibited in young children with ASD can often look

similar to behaviors seen in both typically developing children and those with other early developmental delays (Powell, Heymann, Tsatsanis, & Chawarska, 2018).

Another challenge in early identification is in determining and utilizing optimal practice methods that will promote accurate screening and diagnosis of ASD in young children. At the current time, ASD is diagnosed behaviorally by expert clinicians in medical and educational settings. Although recent advances in neuroimaging and genetic research show promise for future identification of ASD via biomarkers, the current mechanism for diagnosis is ultimately clinical judgment. Ideally, clinical judgment in diagnosis of ASD is supported by reliance on well-validated screening and diagnostic tools, such as the Autism Diagnostic Observation Schedule, Second Editions (ADOS-2; (Lord, Rutter, et al., 2012) or the Autism Diagnostic Interview, Revised (ADI-R; (Rutter, Le Couteur, & Lord, 2003)), to assess for ASD symptoms in a standardized manner. However, there are drawbacks to relying on these assessments, such as long administration times (ranging from 45 minutes to 3 hours), and requirement of specialty training for valid administration and interpretation of results.

Clinical research continues to search for optimized psychometric tools and assessment methods that can improve these issues arising in diagnostic practice. One suggested approach to refining assessment and diagnosis practices is through the utilization of broadband behavioral rating questionnaires as screening tools to be used in the context of diagnostic evaluations with aims to promote efficiency of assessments and facilitate accurate identification (Lord, Corsello, & Grzadzinski, 2014; C. L. Myers,

Gross, & McReynolds, 2014). Broadband behavior rating scales are assessment instruments designed to measure aspects of an individual's behaviors and emotional functioning using a standardized method (Walrath, 2011). For young children, such as those frequently referred for ASD evaluations, behavioral ratings scales are typically completed via caregiver or parent questionnaire, and allow for quick assessment of a wide variety of emotional, behavioral, and adaptive difficulties that can coincide with core symptoms of autism (Bradstreet, Juechter, Kamphaus, Kerns, & Robins, 2017; Havdahl, von Tetzchner, Huerta, Lord, & Bishop, 2016; C. L. Myers et al., 2014)}.

Broadly, behavioral rating scales are routinely administered by clinicians in a variety of settings and are also often included as part of comprehensive psychological evaluations for ASD (Powell et al., 2018). Given the current widespread use of these measures and quick administration time, parent/caregiver-informed questionnaires have the potential to aid in enhancing screening and diagnostic practices within psychological evaluations. Moreover, behavioral rating scales provide rich illustrations of emotional/behavioral profiles that may capture a more comprehensive view of the primary concerns, thereby aiding in diagnostic accuracy (C. L. Myers et al., 2014). Lastly, results of behavioral rating scales have the ability to identify behavioral problems and areas of adaptive difficulty that can help inform treatment recommendations and assist in tailoring early interventions (Bradstreet et al., 2017).

One widely utilized broadband behavioral measure, The Behavior Assessment System for Children (BASC), has been suggested as a tool to aid in screening and

diagnostic practices in ASD populations (Goldin, Matson, Konst, & Adams, 2014; Volker et al., 2010). The BASC is a multimethod, multidimensional rating system that evaluates emotions and behaviors of children and young adults ages 2 through 25 years (Reynolds & Kamphaus, 2015). The third edition, BASC-3, was published in 2015, providing updated normative samples, administration options, and additional items per rating form. The BASC-3 Parent Ratings Scale, Preschool form (BASC-3 PRS-P) contains several clinical and content scales that have been associated with ASD symptomology, such as the Functional Communication, Withdrawal and Atypicality scales, as well as the Developmental Social Disorders (DSD) content scale (Bradstreet et al., 2017; Goldin et al., 2014). Previous studies have evaluated the use of the BASC-2 PRS-P as a screening tool for differentiating young children with ASD from those with other developmental delays and typically developing peers (Bradstreet et al., 2017; Juechter, 2012); however, the performance and utility of the newly updated BASC-3 PRS-P has yet to be examined in a preschool-aged clinical sample.

The current study aims to extend the literature on screening and diagnostic approaches to early identification of ASD in preschool-aged children by investigating the utility of the Developmental Social Disorders (DSD) clinical scale on the BASC-3 PRS-P to identify and distinguish children with ASD among those with other developmental delays or behavioral issues. The study will also investigate the reliability of the DSD scale between English and Spanish questionnaire types, as well as examine potential gender profiles on the DSD scale, as there is growing evidence suggesting girls with ASD

may exhibit more subtle differences in their emotional and behavioral manifestations of ASD symptoms (Matheis, Matson, Hong, & Cervantes, 2018; Sipes, Matson, Worley, & Kozlowski, 2011). Ultimately, these data can help inform practices in the identification of ASD in preschool-age children, as well as provide insight into potential differences in symptom profiles between males and females in this developmental period.

CHAPTER TWO

Review of the Literature

Definition and Prevalence of ASD

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder that affects how an individual thinks, behaves, communicates, and interacts with others. ASD is characterized by impairments in social communication and interactions, along with the presence of restrictive, repetitive behaviors and interests (American Psychiatric Association [APA], 2013). Based on recent estimates from the active surveillance efforts of the Early Autism and Developmental Disabilities Monitoring (Early ADDM) network under the Centers for Disease Control and Prevention (CDC), ASDs are estimated to affect up to 3% of children in the United States (Christensen et al., 2019). Findings from the CDC's report published in April 2019 indicate 1 in 59, 4-year-old children are identified as having an ASD, representing a 20% rise in prevalence for this age group between 2010-2014 (Christensen et al., 2019). These estimates were calculated based on expert review of education and/or health care records for more than 58,000 children across multiple surveillance sites nationwide. Other notable findings in the report include gender ratios ranging from 2.6:1 to 5.2:1 (male: female), as well as a high frequency of 4-year-old children with co-occurring intellectual disability (43.6 to 47% across years) (Christensen et al., 2019). Although the ADDM estimates are not considered to be fully representative of the population, they provide a foundation for understanding the epidemiological characteristics related to ASD and the tracking of the disorder.

Epidemiological research has found that ASD affects individuals across all sociodemographic characteristics; however, prevalence rates vary based on gender, race and ethnicity, and socioeconomic status (Lyall et al., 2017). Regarding gender, ADDM findings document a 4:1 ratio (males: females) for both 4- and 8-year-old groups of boys and girls (Baio et al., 2018), while other studies indicate lower rates ranging from 2:1 to 3.5:1 (Constantino, Zhang, Frazier, Abbacchi, & Law, 2010; Loomes, Hull, & Mandy, 2017).

Prevalence rates are also variable across other sociodemographic factors, such as race and ethnicity. Studies have consistently reported lower rates of ASD diagnoses for Black and Hispanic children as compared to their Caucasian peers (Baio et al., 2018; Dickerson et al., 2017; Mandell et al., 2009), as well as lower prevalence rates in low-income families and those of racial minorities (Dickerson et al., 2017; Durkin et al., 2010; Larsson et al., 2005; Liptak et al., 2008; Nguyen, Krakowiak, Hansen, Hertz-Picciotto, & Angkustsiri, 2016). These disparities in prevalence are thought to be driven by social and public health barriers such as lack of access to diagnostic and treatment services for minority populations, limited knowledge about ASD, language barriers in medical settings, and stigma associated with developmental disorders (Jo et al., 2015).

Onset Patterns

Symptoms of ASD are thought to emerge within the early developmental period and continue to impact development throughout the lifespan. Researchers first began investigating the early behavioral signs of ASD as an approach to inform early

identification and diagnosis and to enhance understanding of the relationships between onset, developmental trajectory, and prognosis (Ozonoff et al., 2011). Initial reports about symptom onset were obtained via retrospective parent reports and thorough review of home videos documenting the behaviors of children who would later receive ASD diagnoses (Goldberg, Thorsen, Osann, & Spence, 2008; Werner & Dawson, 2005; Werner, Dawson, Osterling, & Dinno, 2000). Parent-reported accounts of early atypical development included behaviors such as poor eye contact, extreme temperaments (e.g., persistently calm or persistently irritable), and decreased social responsiveness during the first year of life (De Giacomo & Fombonne, 1998; Werner et al., 2000; Zwaigenbaum et al., 2005). However, these symptoms were not consistently observed across individuals.

Due to the limitations of retrospective studies, other research began to focus on prospectively studying the early signs of ASD in community-based samples and high-risk populations (i.e., infants who have an older sibling with ASD). Findings from longitudinal, prospective studies of high-risk infants found that ASD emerges as early as six months in some cases (Szatmari et al., 2016), but overt behavioral symptoms do not tend to manifest until 12 to 24 months (Jones, Gliga, Bedford, Charman, & Johnson, 2014).

Although the evidence from high-risk infant studies document a subset of children who demonstrate ASD symptoms within the first two years of life, other patterns are sometimes observed. Another onset pattern is characterized by a period of mostly typical developmental progression followed by a loss of previously developed skills (i.e., loss of

language or decreased social engagement) and emergence of atypical behaviors such as repetitive mannerisms and visual examination of objects (Landa, Gross, Stuart, & Faherty, 2013; Ozonoff et al., 2008). An additional identified onset pattern features intact developmental skills that eventually fail to advance or sophisticate, leading to a developmental “plateau,” in which a child’s skills are not progressing as expected (Ozonoff et al., 2018; Shumway et al., 2011). The variability seen in the emergence of ASD symptoms across children likely reflects the heterogeneity also observed within the disorder. While the “early-onset” form of ASD has been suggested as a frequently reported onset pattern, recent studies have indicated the onset pattern characterized by a regression of skills may also be increasingly prevalent (Ozonoff & Iosif, 2019).

Overall, research has determined that initial signs of ASD are present very early in life and commonly first impact social communication skills, with later emergence of restricted, repetitive behaviors. Although some symptoms of ASD may be present in infancy, ASD-related deficits and behaviors may only be detected once a child fails to meet developmental, social, or educational expectations (Baio et al., 2018). Due to the variability in the emergence of ASD symptoms, as well as the rapid behavioral changes that can occur in early childhood, behaviors of ASD are often overlooked or misattributed to other conditions.

Diagnostic Criteria and Clinical Features

Under the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V), Autism Spectrum Disorder is categorized as

a neurodevelopmental disorder characterized by symptoms presenting in two main areas: impairments in social communication and social interaction, and present symptoms of restricted, repetitive behaviors, interests or activities. Additionally, the DSM documents that age of onset of ASD is typically in the early childhood developmental period, although individuals may not be identified as having the disorder until later in life. Lastly, a defining feature of the disorder is that there must be a clinical level of impairment across symptom domains (social impairments and presence of atypical behaviors) that significantly impact a person's daily functioning (APA, 2013).

The first domain, impairments in social communication and interaction, is a core feature of the disorder and is thought to best help distinguish children with ASD from children with other developmental delays and typical development (Sigman, Dijamco, Gratier, & Rozga, 2004). Social communication includes a wide variety of skills, ranging from verbalizations (e.g. spoken language for communicative purposes) (Wetherby, Watt, Morgan, & Shumway, 2007) to nonverbal cues (e.g. eye contact, facial expressions, use of gestures) (Mundy, Sigman, Ungerer, & Sherman, 1986). Individuals with ASD can have a variety of functional language impairments and pragmatic language deficits across these areas. Difficulties in social interaction may include abnormalities in social approach, such as not knowing how to interact with others, or impairments in social reciprocity, including lack of interest in engaging with others and difficulty maintaining the back and forth of a conversation (Constantino, Przybeck, Friesen, & Todd, 2000; S. W. White, Keonig, & Scahill, 2007).

In addition to these social communication and interaction deficits, diagnosis of the disorder requires the presence of at least two of four symptom areas pertaining to restricted, repetitive patterns of behavior, interests, or activities either currently or by history (APA, 2013). These symptoms can include atypical and repetitive motor mannerisms, repetitive use of items, repetitive speech or atypical speech, extreme behavioral rigidity, hyperfocus on specific items or topics of interests, as well as aversions or interests in sensory stimuli (APA, 2013).

Early accounts of ASD reported behavioral observations of children with highly fixated interests in objects and the mechanics of items (Kanner, 1943). Under the current diagnostic model, restricted interests (e.g. an intense interest in trains, interests in sensory stimuli), rigid adherence to routines or rituals (e.g. insisting tasks are performed in a specific order), and repetitive behaviors (e.g., ritualistic play, stereotyped movements) all continue to serve as hallmarks of the disorder (Richler, Bishop, Kleinke, & Lord, 2007). In addition to these core symptoms, motor abnormalities have been heavily documented in subgroups of individuals with ASD (Jansiewicz et al., 2006; Ming, Brimacombe, Chaaban, Zimmerman-Bier, & Wagner, 2008; Mosconi et al., 2015). As mentioned, differences in sensory processing are also highly common within the disorder and it has been suggested that greater than ninety percent of individuals with ASD experience sensory sensitivities, spanning from sensory aversions to sensory interests (Leekam, Nieto, Libby, Wing, & Gould, 2007). Together, deficits in social communication and interaction combined with the presence of restrictive, repetitive behaviors (RRBs), can

produce an array of significant impairments in language, educational and occupational achievement, social skills, and adaptive functioning (Szatmari et al., 2015).

Related Comorbid Conditions. Adding to the complex presentation of ASD symptoms, other clinical features such as co-occurring medical problems, psychiatric conditions, and other neurodevelopmental disorders are also highly prevalent in individuals with ASD across the lifespan (Kuravackel & Ruble, 2014; Lever & Geurts, 2016). Soke, Maenner, Christensen, Kurzius-Spencer, and Schieve (2018a) documented the prevalence of medical and psychiatric comorbidities from ADDM cohorts of 4-year-olds (N=783) and 8-year-olds (N=1091). When combining results from the age groups, over 95% of children had at least one co-occurring condition/symptom; however, 67% of this was attributed to the 8-year-old group. These findings may indicate that comorbidities in young children with ASD are going undetected/undiagnosed, or that symptoms of co-occurring conditions may not manifest until children are older.

Regarding the identification of comorbid conditions in ASD, Mannion and Leader (2013) suggested that a significant barrier in diagnosing comorbidities is due to the absence of validated psychometric instruments that can detect a variety of commonly co-occurring symptoms. Medical comorbidities commonly seen in conjunction with ASD span from gastrointestinal problems, epilepsy, genetic disorders, and cardiac anomalies to feeding disorders (Aldinger, Lane, Veenstra-VanderWeele, & Levitt, 2015; Bauman, 2010; Buie et al., 2010; Doshi-Velez, Ge, & Kohane, 2014; Mari-Bauset, Zazpe, Mari-Sanchis, Llopis-Gonzalez, & Morales-Suarez-Varela, 2014; S. M. Myers & Johnson,

2007). Related and potentially co-occurring neurodevelopmental and genetic disorder can include conditions such as Attention Deficit/Hyperactivity Disorder (ADHD), Intellectual Disability, Down's Syndrome and Fragile X Syndrome, (Moeschler, 2019; Neumeyer et al., 2018; Richards, Jones, Groves, Moss, & Oliver, 2015; Soke et al., 2018a). Other related disorders include psychiatric comorbidities, like anxiety and depression, or other behavioral challenges including self-harming and aggressive behaviors (Hill et al., 2014; Levy et al., 2010; Minshawi, Hurwitz, Morriss, & McDougle, 2015). As any one of these comorbidities can be present alongside ASD and may emerge in early childhood, assessing for the co-occurrence of these conditions proves difficult as preschool-aged children are simultaneously experiencing rapid rates of developmental change.

Unfortunately, the presence of comorbid conditions further complicates achieving an accurate and timely ASD diagnosis. Data collected in 2010 by the ADDM network indicated that age at the time of initial ASD diagnosis both increased *and* decreased when a single co-occurring condition was also documented. Specifically, Soke et al. (2018a) suggest a relationship between the presence of comorbidities in ASD and timing of first diagnosis, such that symptoms of some co-occurring conditions can mask the recognition of core symptoms of ASD, especially in young children. These findings emphasize the importance of clear and valid early identification practices as clinicians are required to distinguish peripheral symptomology from core symptoms of ASD in a timely and accurate manner. In line with this, further research is needed to investigate optimal ways that clinicians can evaluate symptoms of related conditions and behaviors

associated with ASD in order to determine the source of a child's behavioral presentation and identify appropriate diagnoses and treatments.

Considerations Related to Gender. An area of research that has gained momentum in recent years is the differences in manifestation of ASD symptoms between boys and girls. However, research investigating differences in ASD symptoms across gender in the preschool-aged population has yielded mixed findings. Some previous studies have documented no significant differences between young boys and girls with ASD (Carter et al., 2007; Holtmann, Bölte, & Poustka, 2007; Reinhardt, Wetherby, Schatschneider, & Lord, 2015), while others found evidence for more significant social communication deficits in young girls with ASD (Hartley & Sikora, 2009; Lawson, Joshi, Barbaro, & Dissanayake, 2018) along with reduced rates of restrictive, repetitive behaviors and interests (Sipes et al., 2011). Interestingly, this feature of reduced RRBs was found to be especially true for preschool-aged girls with ASD who had higher cognitive abilities and language skills (i.e., $IQ > 70$) (Giarelli et al., 2010; Knutsen, Crossman, Perrin, Shui, & Kuhlthau, 2019). Given that there is a lack of gender-specific diagnostic instruments to help clinicians distinguish behavioral differences, gender differences in behavioral manifestations of ASD symptoms may make it more difficult for clinicians to accurately identify the disorder in young children.

Early Identification of ASD

The process of early identification of ASD can be broken down into three broad stages: 1) detection of developmental concerns by parent/caregiver or pediatrician, 2)

confirmation of developmental delays via screening/developmental surveillance practices, followed by referral for further assessment, and 3) completion of a comprehensive diagnostic evaluation to determine an appropriate diagnosis (Lappe et al., 2018). Currently, ASD is detected and confirmed through means of behavioral observations and use of psychometric tools, as there are no biologically based diagnostic tests that can confirm its presence. Still, opportunities for early identification of ASD have significantly improved in recent years through an expansion of scientific knowledge and resulting community awareness. The CDC's "Learn the Signs. Act Early" campaign and similar efforts from Autism Speaks® use widely circulated media, such as billboards, television ads, and smart phone applications, to help parents and caregivers note potential behavioral signs and developmental concerns early (CDC, 2019).

An additional area of focus has been placed on education and outreach for pediatricians to improve screening and use of diagnostic tools to make initial identification of children at risk for an ASD. An accumulation of evidence documenting the stability of ASD diagnoses in children as young as 18 months (Ozonoff et al., 2015; Zwaigenbaum et al., 2016) has further supported the focus on early identification and places pediatricians at a particularly crucial position to serve in this capacity. Thus, several recommendations have been made for this purpose, including identifying early markers of ASD that can be measured in routine clinical practice settings, such as the pediatrician's office, as well as the suggestion to implement universal screening practices (i.e., screening all children at periodic well-child appointments regardless of if there is a

clinical concern) (Khowaja, Robins, & Adamson, 2018). Finally, guidance on improving referral methods for a diagnostic evaluation after a positive screen for ASD (Gordon-Lipkin, Foster, & Peacock, 2016) has also been provided.

Following initial identification, best-practice guidelines for the diagnosis of ASD in young children recommend a comprehensive, multidisciplinary evaluation that employs a team of medical and clinical professionals to assess a child's functioning across a variety of developmental domains (Johnson & Myers, 2007; S. M. Myers & Johnson, 2007; Zwaigenbaum et al., 2009). This model often includes formal, standardized assessments of ASD symptomology, as well as assessment of a child's current cognitive level, physical functioning, speech/language abilities, and emotional, social, and adaptive areas of functioning. Practice parameters also indicate the need for direct behavioral observation and a thorough review of medical, developmental, and educational records in order to provide context to a child's behavioral presentation and inform differential diagnosis (Volkmar et al., 2014). Furthermore, guidelines promote the importance of obtaining information about a child's current behaviors from multiple sources (i.e., parents/guardians, teachers), across multiple settings in order to accurately interpret pervasive behavioral patterns (Ozonoff, Goodlin-Jones, & Solomon, 2005).

Commonly used "gold standard" diagnostic tools for ASD include standardized observational measures, such as the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) and the Childhood Autism Rating Scales, Second Edition (CARS-2), which allow clinicians to create play-based scenarios with a child aimed at eliciting

symptomology related to ASD (Lord, Rutter, et al., 2012; Schopler, Van Bourgondien, Wellman, & Love, 2010; Zwaigenbaum et al., 2009). Parent-reported behavioral questionnaires, such as the Social Communication Questionnaire (SCQ), briefly assess for the presence of specific ASD symptoms, including the intensity, frequency and duration of these behaviors, based on parental knowledge of the child's overall behaviors (Barnard-Brak, Brewer, Chesnut, Richman, & Schaeffer, 2016; Rutter, Bailey, & Lord, 2003). More extensive interview-based autism diagnostic tools, such as the Autism Diagnostic Interview-Revised (ADI-R), allow for collection of rich information of a child's developmental history in conjunction with assessing for nuances of ASD symptomology (Lord, Rutter, & Le Couteur, 1994; Soke et al., 2011). The complete evaluation process is extensive and, together with the increasing prevalence, results in lengthy wait times for families and taxing demands for clinicians. Clearly, there is great need for streamlined evaluations and diagnostic instruments with enhanced specificity in order to allow early identification of ASD and access to evidence-based interventions that have been shown to improve long-term developmental trajectories (Zwaigenbaum et al., 2015b).

Use of Emotional/Behavioral Rating Scales in the Identification of ASD.

In line with recommendations for further innovations in diagnostic instruments, one method to enhance screening and diagnostic accuracy and efficiency is through the use of broadband behavioral rating scales (Khowaja et al., 2018; C. L. Myers, Bour,

Sidebottom, Murphy, & Hakman, 2010; C. L. Myers et al., 2014; S. P. White et al., 2012; Zwaigenbaum et al., 2015a; Zwaigenbaum et al., 2015c; Zwaigenbaum et al., 2009).

Broadband rating scales are questionnaires commonly given to parents/caregivers to measure aspects of their child's emotional and behavioral functioning, such as self-regulation, emotional reactivity, anxiety, disruptive behaviors, inattention/hyperactivity, and sleep disturbances (C. L. Myers et al., 2014; Powell et al., 2018; Zwaigenbaum et al., 2015c). While these questionnaires are often used as a part of the comprehensive diagnostic evaluation for ASD, several studies have promoted use of these measures to further aide in differential diagnosis by helping differentiate between core ASD behaviors versus symptoms related to another disorder, as well as to help screen for behaviors indicative of a comorbid medical or psychiatric condition (Bauman, 2010; Chawarska, Klin, & Volkmar, 2008; King, 2016; Landa et al., 2013; Lever & Geurts, 2016; Stone, Ousley, & Littleford, 1997; Volkmar et al., 2014; Zwaigenbaum, Bryson, & Garon, 2013).

Two popular broadband rating scales, the BASC-3 and the CBCL, are often administered to screen for symptoms of psychopathology in children and adolescents and are commonly utilized across school, medical and forensic settings (Merrell, 2008; Shapiro & Heick, 2004). Due to the well-established validity of these measures and their widespread use across settings, behavioral rating scales have been suggested as a potential alternative method for developmental screening for ASD (Huerta & Lord, 2012; C. L. Myers et al., 2014). Moreover, these measures are efficient in that they are

relatively quick to complete (i.e., around 30 minutes total) and do not require extensive clinical training to administer or score. Both the CBCL and BASC parent-reported questionnaires contain items that assess for behaviors and symptoms that align with ASD, and thus, a few studies have examined the validity of these measures in ASD populations.

Research into the clinical utility of behavioral ratings scales in ASD populations have supported use of both the CBCL and BASC parent rating scale instruments as effective screening tools to identify concerns for ASD in general pediatric offices, schools, and other hospital-based settings (Bradstreet et al., 2017; Gardner, Campbell, Bush, & Murphy, 2018; Juechter, 2012; Muratori et al., 2011; C. L. Myers et al., 2014; A. Narzisi et al., 2013). These scales have also been shown to be useful in assessing and monitoring response to ASD interventions (Guli, Semrud-Clikeman, Lerner, & Britton, 2013; Leaf et al., 2009; Thomson, Burnham Riosa, & Weiss, 2015). Though both scales have been looked at, historically, the CBCL has been the primary behavioral rating scale used for screening and assessment due to the extensive amount of research studies delineating the utility of two of its clinical subscales, the Pervasive Developmental Disorders (PDD) scale and Withdrawn scale, to distinguish school-aged children and adolescents with ASD from those who are typically developing (Havdahl et al., 2016; Muratori et al., 2011; Rescorla et al., 2017). Further investigation of the BASC is necessary.

Use of the BASC Parent Rating Scales in ASD Populations.

Though the use of the second edition of the BASC, (BASC-2) has been validated in various ASD age-groups, particularly in the use of its Developmental Social Disorders (DSD) scale, Withdrawn scale, and Adaptability scales (Bradstreet et al., 2017; Hass, Brown, Brady, & Johnson, 2012; Kent, 2006; Mahan & Matson, 2011; Volker et al., 2010), past research has been limited. In general, studies have examined the functionality and use of BASC-2 Parent Rating Scales as a screening instrument to distinguish between children and adolescents with ASD versus those who are either a) typically developing or b) have another developmental delay or psychiatric disorder. First, findings from studies comparing BASC-2 scores in ASD groups and typically developing (TD) groups of children ages 6-16 reported that children with ASD obtained significantly higher average scores on scales that assess symptomology related to hyperactivity, depression, anxiety, hyperactivity, and atypical development than controls, and that scores on the DSD scale could predict group membership when the cutoff score for clinical significance was set at 60 (Mahan & Matson, 2011; Volker et al., 2010).

For findings from studies that employed the BASC-2 scales to differentiate between ASD groups, TD groups, *and* a developmental delay (DD) group (i.e., children with other developmental disorders separate from autism) in a children ages 2-16 found that children with ASD scored significantly higher than both the TD and DD groups on the Developmental Social Disorders (DSD) scale, the social withdrawal scale and atypicality scale (Goldin et al., 2014). They also noted that the ASD group scored significantly below the DD and TD groups on various scales that measured different

aspects of adaptive functioning, such as social skills, leadership, activities of daily living, and functional communication. The BASC system primarily intended the DSD scale, to assess for a clinical level of symptoms related to developmental social differences in children. The authors' current definition for the DSD scale is a scale that examines "the tendency to display behaviors characterized by deficits in social skills, communication, interests, and activities; such behaviors may include self-stimulation, withdrawal, and inappropriate socialization" (Reynolds & Kamphaus, 2015).

A handful of studies have examined the functionality and clinical utility of the DSD scale from the second edition of the BASC, the BASC-2, in preschool-aged children. Juechter (2012) compared BASC-2 PRS-P DSD scores and overall profiles for preschoolers ages 25 to 37 months divided into three groups (children with ASD or PDD-NOS, other developmental delays, and typical development). Results from analyses showed that preschoolers with ASD had higher mean T-scores on the DSD scale than those children who were given diagnoses of PDD-NOS ($M=69.79$ vs. $M=61.11$, respectively). Interestingly, they found no significant differences between mean DSD scores when comparing the combined ASD and PDD-NOS group to the other developmental delays group, meaning that the DSD scale was not specific enough to differentiate between those preschoolers with ASDs and those with other developmental disorders. However, the DSD scale was found to have adequate sensitivity in distinguishing between all clinically referred children (ASD, PDD-NOS, and Other developmental disorders) and typically developing children. Furthermore, Juechter

(2012) compared clinical and adaptive profiles among the combined ASD and other developmental delays group and found no statistically significant differences.

In contrast to these findings, C. L. Myers et al. (2014) evaluated the clinical and adaptive scales of the BASC-2 PRS-P as a possible screening instrument for ASD in a sample of clinically referred preschoolers (ages 24-71 months) and found that the ASD group obtained significantly *lower* scores than the group with other diagnoses or developmental delays on clinical scales for Hyperactivity, Aggression, Anxiety, and Depression. However, the ASD group scored significantly higher on scales that measured difficulties with adaptive skills, including scales for Social Skills, Functional Communication. These findings seem to suggest that perhaps preschoolers with ASD may display poorer adaptive skills, differentiating them from children with developmental concerns but no ASD. In addition, Myers (2014) did not examine the DSD scale as part of this study.

Bradstreet et al. (2017) expanded on studies conducted by Juechter (2012) and Myers (2014) by examining the BASC-2 PRS-P DSD scale, clinical, and adaptive profiles among two groups: children with ASD and a “non-ASD” group that was comprised of children with other developmental delays, other diagnoses (e.g., ADHD, epilepsy) and typically developing children. Furthermore, they assessed the concurrent validity of the DSD scale from the BASC-2 PRS-P and other prominent ASD screening and diagnostic tools, including the Modified Checklist for Autism in Toddlers, Revised (M-CHAT-R; (Robins et al., 2014)), the Childhood Autism Rating Scale (first and second

editions; (Schopler, Reichler, DeVellis, & Daly, 1980; Schopler et al., 2010)), and the ADOS (first and second editions; (Lord, Rutter, et al., 2012)). Results from analyses determining the clinical utility of the DSD scale indicated that the DSD scale was able to accurately differentiate 72 % of preschoolers with ASD and 63 % of preschoolers without an ASD diagnosis; however, these percentages are influenced by the composition of the various diagnoses and non-diagnoses of the comparison group. They did note that the DSD scale was better at distinguishing between those with ASD, other developmental delays, and other diagnoses versus children who were typically developing, but reported limited ability for the DSD scale when comparing clinical subgroups. Lastly, Bradstreet et al. (2017) reported moderate positive correlations between the DSD scale and the M-CHAT-R, CARS-2, and ADOS-2, establishing adequate concurrent validity among these measures.

Finally, Gardner, Campbell, Bush, and Murphy (2017) examined BASC-2 PRS-P DSD scores among children (ages 24 to 71 months) with ASD (ASD), children with comorbid ASD and ID (ASD/ID), and children with ID (ID), looking at comparisons among two specific subgroups of neurodevelopmental disorders. Additionally, the study aimed to examine whether any group differences on the DSD, clinical, and adaptive scales emerged between African American and Caucasian racial groups. Results of the study specified that the BASC-2 DSD scale produced equally elevated scores for preschoolers diagnosed with ASD (both with and without comorbid ID) when compared with preschoolers diagnosed with ID only, indicating poorer specificity of the DSD scale

when applied to clinical samples. Moreover, they found no significant differences between racial groups on BASC-2 clinical and adaptive scales, and comparable accuracy of the DSD scale in screening for ASD vs non-ASD.

Changes and updates from the BASC-2 to the BASC-3 PRS-P. In 2015, authors of the BASC released an updated version, the Behavior Assessment System for Children, Third Edition (BASC-3) which included several significant changes. Prior to discussing changes, it should be noted that all items from the BASC-2 PRS forms were included in the development and standardization of the BASC-3 PRS forms; however, additional items were added within the development of the BASC-3 edition as a result of the field studies for standardization (Reynolds & Kamphaus, 2015). It should also be mentioned that BASC-2 PRS and BASC-3 PRS scores can be computed and compared using the PRS T-Score Mean Differences table provided in the BASC-3 manual (Reynolds & Kamphaus, 2015).

Specifically focusing on updates from the BASC-2 and BASC-3 parent rating forms, the BASC-3 PRS forms now include enhanced clinician usability in terms of digital scoring and online reporting options, updated and expanded norm/comparison groups (for both general and clinical normative samples), as well as separate normative samples for scoring Spanish-language PRS forms (Reynolds & Kamphaus, 2015). Spanish versions of the BASC-2 were available for use but the BASC system did not develop separate normative samples for Spanish-language individuals, thereby making the interpretability of the Spanish measures questionable. In the updated BASC-3,

separate normative samples were developed and validated for both general and clinical norms for Spanish-language BASC-3 PRS versions (allowing for enhanced accuracy of scoring); however, there have been very few follow-up studies that have confirmed the validity of use of the Spanish norms for clinical purposes.

Additionally, the normative groups for the BASC-3 PRS were also broken down by gender and sub-age groups to allow for more choices in selecting an appropriate comparison group relevant to the child's presenting clinical concern. The BASC-3 PRS now offers a choice of either general or clinical norm sample; general norms are representative of the general population, while clinical norms represent scores from children who have been diagnosed with a variety of behavioral and emotional disorders (Reynolds & Kamphaus, 2015). Scorers of the BASC PRS can also select among gender normative samples, which include sample for "combined-gender (i.e., male and female) or separate-gender (i.e., only females or only males). Authors of the BASC argue that allowing scores to be calculated based on specific genders or by clinical features allows clinicians to better answer a variety of clinical questions (Reynolds & Kamphaus, 2015).

Looking more specifically at the updates made to the BASC parent rating scales for the preschool-aged population (ages 2-5), the main differences between the BASC-2 PRS-P and BASC-3 PRS-P include updates to the general and clinical normative comparison groups and the expansion of items across several clinical scales, including the Developmental Social Disorders (DSD) scale. Interestingly, only children ages 4-5 years-old were included in the development of the clinical normative samples for the

preschool age group, despite the age-range for the BASC-3 PRS-P including both 2- and 3-year-olds. Authors cited that clinical diagnoses in the 2- and 3-year-old ages are less reliable (Reynolds & Kamphaus, 2015). In terms of additional items added to the DSD scale of the BASC-3 PRS-P, two items were added that asked parent-responders to rate the frequency that their child “Avoids eye contact” and “Engages in repetitive movements”(Reynolds & Kamphaus, 2015).

Considering these substantial updates in test content and normative comparison groups between the BASC-2 and BASC-3 PRS versions, future studies are needed to investigate the clinical impact of these changes on diagnostic and intervention practices in children. Perhaps more importantly, the updates included within the BASC-3 PRS-P now offer an opportunity to re-examine this emotional/ behavioral rating scale in ways that were not previously feasible (i.e., examining scores between Spanish/English versions, and examining scores based on different choices of gender norm groups). For ASD and developmentally delayed populations of preschool-aged children, gaining a better understanding of the utility of the updated BASC-3 parent rating scale could potentially enhance diagnostic and early identification practice, as well as provide important behavioral information that could be used to inform plans for individualized intervention services.

Summary and Aims of the Current Study

Early screening and identification of ASD is important to ensure initiation of interventions that have been shown to positively impact developmental trajectories in

young children with ASD (Zwaigenbaum et al., 2009). In order to enhance optimal rates of early diagnosis, screening and diagnostic assessment techniques must continue to improve. Current practice recommendations encourage the use of a comprehensive assessment model to identify ASD in preschool-aged children. Although a comprehensive approach to ASD evaluations facilitates diagnostic accuracy, criticisms include ongoing difficulties in differential diagnosis when assessing young children with developmental delays and lengthiness of the assessments, which contributed to families' stress, delays diagnoses and postpones access to effective treatments (Lappé et al., 2018; C. L. Myers et al., 2014).

One suggested method for enhancing screening and diagnostic accuracy is through the use of broadband behavioral rating scales (Havdahl et al., 2016; Lord et al., 2014; C. L. Myers et al., 2014) as they have the ability to assess a wide variety of psychopathology, including symptoms of ASD and developmental disorders (Reynolds & Kamphaus, 2015), are commonly administered to children across multiple settings (e.g., school, pediatrician offices, specialty clinics), and are quick and easy for clinicians to administer. Previous research investigating use of behavioral ratings scales to aid in detection of ASD in early childhood populations highlights the utility of measures such as the CBCL and BASC to differentiate between ASD and typical development; however, these studies have reported low evidence that these rating scales can distinguish ASD from other developmental delays in clinically-referred preschool samples.

Moreover, use of the CBCL in preschool populations is well-established, while use of the BASC in this population is comparatively underdeveloped.

The current study aims to extend previous research on use of the BASC PRS in preschool-aged populations by exploring the diagnostic and clinical utility of the Developmental Social Disorders (DSD) scale on the updated version of the BASC (third edition; BASC-3 PRS-P) to differentiate between young children with autism spectrum disorder and those with other developmental delays referred for a developmental diagnostic evaluation. Furthermore, the study sought to investigate the reliability of the DSD scale to detect ASD when comparing BASC-3 PRS-P questionnaire language types (Spanish vs. English), and differences in DSD scores for girls and boys within a clinically-referred sample.

Research Questions and Hypotheses

The current research extends and expands previous BASC-2 PRS studies in preschool-aged populations by addressing the following questions/sub-questions:

1. Is the BASC-3, Developmental Social Disorders (DSD) scale a reliable scale in the present clinical sample? As part of this broader question:
 - a. Do scores obtained on the DSD scale demonstrate concurrent validity with scores obtained the ADOS-2, a “gold standard” diagnostic tool in the detection of symptoms of ASD (Powell et al., 2018)?
 - b. Are there significant differences in reliability of the DSD scale between English and Spanish versions of the BASC-3 PRS-P?

2. Can the BASC-3, DSD scale be used as a tool to reliably identify and differentiate children with ASD and those with other developmental delays in the present sample? As part of this broader question:
- a. Are there significant group differences (ASD vs. Non-ASD) in DSD scores when scores are calculated using three different normative comparison groups? : a General normative group, a Clinical normative group, and a Gender-Specific Clinical normative group.
 - b. Which statistically derived cut-score or range of cut-scores on the DSD scale best differentiates children with ASD from those with other developmental delays?
 - c. Based on a selected DSD cut score, is this cut score equally valid for boys and girls within this sample?

Given findings from previous research using the BASC-2 PRS-P, we hypothesized that the DSD scale would demonstrate adequate concurrent validity with scores from the ADOS-2, a widely used diagnostic tool in the assessment of ASD. Further we hypothesized that there would not be significant differences in reliability of the DSD scale when comparing scores from Spanish and English versions of the BASC-3, as adequate internal consistency values were reported in the BASC-3 manual. In terms of discriminative ability of the DSD scale, we hypothesized that the scale would not be able to reliably differentiate between the ASD and non-ASD groups. This hypothesis was primarily based on the results from studies conducted by Juechter (2012), Bradstreet et al.

(2017), and Gardner et al. (2017) which indicated that the BASC-2 DSD scale is best utilized as a screener within the general population as a way to distinguish children with developmental delays from those who are typically developing, rather than differentiating between clinical populations. A cut score (or range of cut scores) will be determined based on weighing aspects of clinical utility in terms of sensitivity and specificity values. Lastly, we hypothesized that the statistically derived cut score will be equally valid for boys and girls within the sample given that behaviors related to ASD have been reported to be clinically similar across genders during the early developmental period (Powell et al., 2018).

CHAPTER THREE

Methodology

Sample Recruitment

The study implemented a retrospective, unmatched case-control study design, comparing scores obtained on the DSD scale on the BASC-3 PRS-P questionnaire in preschool-aged children with ASD and those with other developmental delays. The study was conducted at the Center for Autism and Developmental Disabilities (CADD) at UT Southwestern Medical Center and Children's Health in Dallas, Texas. One hundred and fifteen participants were selected for study participation from a clinical database (EPIC software) of patients who were seen at CADD for a diagnostic evaluation.

Formal inclusion criteria for the study included that 1) the child was seen at CADD for a diagnostic evaluation and was diagnosed with an Autism Spectrum Disorder or another developmental delay, 2) the child was between the ages of 2-5 years old at the time of the evaluation 3) the child's parent/caregiver completed a BASC-3 PRS-P as part of the child's psychological evaluation and clinical care. Children with known genetic disorders, physical disabilities, and neurological disorders were included in this study, to acknowledge that children with ASD may have co-occurring medical conditions. Both English and Spanish versions of the BASC-3 PRS-P forms were included. Exclusion criteria for this study included children outside the ages of 2-5 years old who were seen for a diagnostic evaluation at the CADD, and children who were between the

ages of 2-5 years old at the time of their diagnostic evaluation but did not receive a BASC-3 PRS-P rating scale as part of their overall assessment battery.

Procedures Utilized for Research

Caregivers of study participants completed the BASC-3 PRS-P behavioral rating scale as part of their child's comprehensive psychological evaluation at CADD. The full clinical evaluation included administration of developmental testing, assessment of adaptive functioning, administration of a play-based/observational assessment, as well as a parent/caregiver interview and review of educational and medical records. Evaluations were conducted by a licensed clinical psychologist specializing in the assessment of ASD, along with other members of the assessment team (i.e., Speech-Language Pathologist, trained graduate students, psychometricians). Final diagnoses were determined by the psychologist in accordance with DSM-V diagnostic criteria.

A waiver for informed consent was obtained through the Institutional Review Board (IRB) at UT Southwestern, as the study was retrospective in nature and data to be examined were included as part of the standard of care. Potential participants were identified through two electronic data sources: 1) Children's Health EPIC database for those patients seen at the Center for Autism and Developmental Disabilities (CADD) for a diagnostic evaluation, and 2) Q-Global: an electronic scoring software that is used to record and score BASC-3 questionnaires. Both systems were accessed by approved study personnel using Children's Health secured computers.

As part of the participant ascertainment process, we first use the EPIC system to conduct a query that identifies patients who received a developmental, diagnostic evaluation between the dates of 01/01/2016-10/18/2019 and who were between the ages of 2-5. From this report, we temporarily recorded and stored selected items of the patient's identifying and demographic information, including name, Medical Record Number (MRN), date of birth, date of evaluation, gender, race, and ethnicity, in a Microsoft Excel Spreadsheet, which was password protected and saved on a Children's Health secure network server. Only the primary investigator had access to this information.

Following this, we utilized the Q-Global software system to query a list of completed BASC-3 PRS-P questionnaires. Then, examining the names and MRN numbers populated from both the EPIC and Q-global, we identified which patients were seen for a diagnostic evaluation and whose caregivers also completed a BASC-3 PRS-P. This process produced a group of participants whose data were reviewed and abstracted for further investigation and analyses. Upon final identification of study participants, the names of all patients (included and excluded in the study), as well as all the abstracted information from excluded patients who were originally identified via EPIC, were deleted. A unique study identification number (UID) was assigned to each subject's MRN. The MRNs and UIDs were stored in a separate, password protected Microsoft Excel document that was housed on a computer operating under Children's Health secured network. This information was maintained by the study investigator for

the duration of the study and was deleted upon study completion. Subject's UUIDs, dates of birth, dates of evaluations, genders, races, and ethnicities were stored in another Excel document, using password protection, and was saved on the Children's Health secured network system.

Finally, we examined and recorded scores from other clinical measures that were obtained at the time of the participants' diagnostic evaluations, including the type of standardized developmental/cognitive assessment that was administered (i.e., the Mullen Scales of Early Learning [MSEL], the Differential Ability Scales, Second Edition [(DAS-II)], or the Leiter Performance Scale, Third Edition [Leiter-3]), the verbal and nonverbal composite scores and full scale IQ scores from these cognitive measures, as well as the algorithm scores from the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2). Other autism-specific behavioral rating measures were collected for diagnostic purposes during the time of the evaluation but were not examined as part of this study. Data stored in Excel were transformed and transferred to Statistical Package for the Social Sciences (SPSS) software for statistical analyses.

Measures

Behavior Assessment System for Children, Third Edition (BASC-3). The BASC-3 is a comprehensive, multidimensional, multimethod behavior rating system that assesses behaviors, emotions, and self-perceptions of children, adolescents, and young adults (Reynolds & Kamphaus, 2015). The BASC-3 contains several forms, including the Parent Rating Scales (PRS), Teacher Rating Scales (TRS), and Self-Report of Personality

(SRP). As the Parent Rating Scale-Preschool Form (PRS-P) is the only form used in this study, a detailed description is provided below.

The BASC-3 PRS-P was published in 2015. It is available in both English and Spanish versions and is suitable for children ages 2-5. Updates from the BASC-2 PRS-P included updated general and clinical norms, as well as a few edits to specific items (i.e., additions, deletions). The BASC-3 PRS-P parent/caregiver questionnaire is comprised of 138 items that require the respondent to rate each item using a 4-point Likert scale (i.e., 0=Never, 1=Sometimes, 2=Often, and 3=Almost Always) assessing the frequency the child displays the behavior in question.

Overall, the BASC-3 PRS-P assesses both broad and narrow behavioral domains, as well as maladaptive and adaptive behaviors. Item responses are summed and converted into standardized T-scores with a mean of 50 and a standard deviation of 10. Higher T-scores on the clinical scales are thought to indicate a higher frequency of emotional or behavioral problems. T-scores ranging from 60-69 are considered “at-risk” while T-scores of 70 or greater are considered “clinically significant.” In contrast to this, lower scores on the adaptive scales are suggestive of adaptive deficits, with T-scores ranging from 31-40 in the “at-risk” range and scores equal to or less than 30 indicated as “clinically significant.” The BASC-3 PRS-P reports T-scores for behaviors on the following scales: Hyperactivity, Aggression, Anxiety, Depression, Somatization, Attention Problems Atypicality, Withdrawal, Adaptability, Social Skills, Activities of Daily Living, and Functional Communication. Additionally, T-scores are reported for

seven content scales: Anger Control, Bullying, Developmental Social Disorders, Emotional Self-Control, Executive Functioning, Negative Emotionality, and Resiliency; four composite scales: Adaptive Skills, Externalizing Problems Index, Internalizing Problems Index, and the Behavioral Symptoms Index; and two clinical indices: Clinical Probability Index and the Functional Impairment Index.

Standardization of the BASC-3 PRS-P was derived from an overall sample of 683 children ages 2-5. Two norm groups were constructed: the general norm sample and clinical norm sample. From the total sample, the general norms are composed of data obtained from a representative sample of children across the United States. The authors attempted to resemble the population with a relative representation of genders, socioeconomic status, race/ethnicity, and classification in special education or gifted/talented programs. The clinical norm sample within the BASC-3 PRS-P was comprised of 83 children (8% of the total sample) between the ages of 4-5 years old identified as having a clinical diagnosis or educational classification of one or more emotional or behavioral problems. Data from children ages 2-3 were not used in the standardization of the clinical norm sample. Furthermore, in the clinical sample, approximately 40 % of the children were female while 60% were male, and 48% of children were White, Non-Hispanic.

Psychometrically, the authors reported adequate reliability and validity for both English and Spanish versions of the BASC-3 PRS-P. Relevant to the research questions of the current study, reliability for the Developmental Social Disorders (DSD) scale was

evidenced by the following: acceptable to excellent reliability coefficients (Cronbach's α =.80-.90 and α =.87 for the English and Spanish versions respectively) across general and clinical samples. For the English version, the DSD scale showed moderate test-retest reliability ($r = 0.89$), and adequate interrater reliability of ($r = 0.79$). In terms of evidence of validity for the DSD scale, intercorrelations revealed positive correlations with scales and indices that would be expectedly elevated, such as atypicality ($r=.71$), executive functioning ($r=.79$), and functional impairment ($r=.92$), while negative correlations were observed for scales such as adaptive skills ($r=-.81$), functional communication ($r=-.76$), and resiliency ($r=-.81$). The DSD scale on the BASC-3 PRS-P is also shown to be moderately correlated with the DSD scale on the BASC-2 PRS-P ($r=.86$) and demonstrates adequate correlations with the Autism Spectrum Rating Scales (ASRS) Total Score ($r=.63$) and the ASRS DSM-IV-TR Score ($r=.66$). Lastly, the DSD scale is adequately correlated with the Achenbach System of Empirically Based Assessment, Child Behavioral Checklist, Ages 1.5-5 Total Problems score ($r = 0.63$) and poorly correlated with the Pervasive Development Problems score ($r=.57$).

Mullen Scales of Early Learning (MSEL). The MSEL is an instrument that assesses early intellectual and motor development, as well as school readiness and response to intervention (Mullen, 1995). It can be administered to children ages 0 to 5 years, 8 months (68 months) to assess learning styles, strengths, and weaknesses, as well as determine the need for special services. Use of the MSEL in the assessment of developmental level of functioning for young children with autism is common both

clinically and in research as its administration is flexible and can be adapted to meet behavioral challenges often displayed by children with ASD (Akshoomoff, 2006; Bishop, Guthrie, Coffing, & Lord, 2011; Bishop, Luyster, Richler, & Lord, 2008; Powell et al., 2018; Stephens et al., 2018).

The MSEL provides T-scores ($M = 50$; $SD = 10$), percentile ranks, and age equivalents that can be calculated across five emerging areas of cognitive and motor functioning: Gross Motor, Fine Motor skills, Visual Reception (visual problem-solving) skills, and Expressive and Receptive language skills. A total composite score, the Early Learning Composite (ELC) is comprised of scores obtained by the four core subtests, visual reception, fine motor, expressive, and receptive language. The authors of the MSEL report overall adequacy of psychometric properties. Internal consistency ranged from .75 to .83 amongst individual scales, and .91 for the ELC; test-retest reliability coefficients were not available for assessment during original standardization. Interrater reliability was noted to be from .91 to .99 when examining children ages 1 - 44 months. Concurrent validity of the MSEL ELC and Bayley Mental Developmental Index (MDI) was .53 – .59. The MSEL gross motor scale was highly correlated with the Bayley Psychomotor Development Index (.76); however, a weak correlation was found for the MSEL gross motor with the Mental Development Index of the Bayley (.30) (Burns, King, & Spencer, 2013).

As noted above, the composite total score is comprised of a child's abilities in the areas of Visual Reception, Fine Motor, Expressive Language, and Receptive Language;

however, as difficulty completing all subtests is common when testing young children with ASD who exhibit behavioral problems (i.e., noncompliance, aggressive behaviors, etc.), the MSEL also includes a Nonverbal Developmental Quotient (NVDQ) that can be obtained with only the Visual Reception scale and Fine Motor scale age-equivalents (Akshoomoff, 2006; Bishop et al., 2011; Stephens et al., 2018). This value is transformed to a standard score than can be compared to other nonverbal IQ measurements (Stephens et al., 2018). The calculation of the NVDQ, as cited by Stephens et al. (2018) is as follows:

$$\text{NVDQ} = \frac{\text{average (of VRae and FMae)}}{\text{chronological age (months)}} \times 100$$

In the present study we computed NVDQ scores from the Mullen Scales of Early Learning in order to compare and potentially combine these scores with values from other measures of nonverbal IQ on the DAS-II and Leiter-3. Bishop et al. (2011) conducted an important study examining the convergent validity of scores from the MSEL and scores from the DAS-II, which yielded results that suggested “good convergent validity with respect to nonverbal IQ (NVIQ), verbal IQ (VIQ), and NVIQ–VIQ profiles. These findings provide preliminary support for the practice of using MSEL age-equivalents to generate NVIQ and VIQ scores.” The concurrent validity of the MSEL nonverbal developmental quotient and various other nonverbal IQ measures has since been validated in several other studies, thus confirming accuracy and utility of its use

(Akshoomoff, 2006; Bishop, Farmer, & Thurm, 2015; Farmer, Golden, & Thurm, 2016; Stephens et al., 2018)

Differential Ability Scales, Second Edition (DAS-II). The DAS-II is a test of cognitive abilities that can be used with children and adolescents ages 2 to 17 (Elliott, 2007). Relevant to the current study, the Early Years battery of the DAS-II is specific to the preschool age range and is appropriate to administer to children between ages 2 years, 6 months and 6 years, 11 months. This specific battery is broken into 2 levels, the Lower Level (ages 2:6 - 3:5) and Upper Level (ages 3:6 - 6:11). The Lower level has four core subtests that assess verbal and nonverbal abilities, including measures of auditory and visual working memory, as well as numerical concepts. The four core subtests contribute to a general conceptual ability score (GCA), commonly used at the total composite score. Similarly, the Upper level consists of 6 subtests that broadly assess verbal skills, nonverbal reasoning, and spatial abilities. Scores on these subtests ultimately aggregate to a GCA. For the Upper level, there are also 11 additional subtests that can be given to assess areas of school readiness, processing speed, and working memory.

Authors of the DAS-II report the following reliability characteristics for the Early Years subtests: internal reliability scores and cluster/composite scores range from .79 to .94, test-retest reliability coefficients (retest between 7-63 days) ranging from .51 to .92, interrater reliability of .98 to .99 in the normative sample. In terms of validity, concurrent validity was shown to be moderate to high with the Wechsler Preschool and Primary Scale of Intelligence–Third Edition and the Wechsler Intelligence Scale for Children–

Fourth Edition. Important for this study, the DAS-II has been shown to have high convergent validity with the Mullen Scales of Early Learning (MSEL) in a sample of children with autism and other developmental disabilities (Bishop et al., 2011). Measures of nonverbal IQ (NVIQ) and verbal IQ (VIQ) between the DAS-II and MSEL were adequate-moderately correlated (.74; and .82, respectively).

Leiter Performance Scale, Third Edition (Leiter-3). The Leiter-3 (Roid, Miller, Pomplun, & Koch, 2013) is a cognitive assessment tool designed to evaluate nonverbal intellectual ability, memory, and attention in individuals ages 3 years to greater than 75 years-old. The administration of the Leiter-3 does not require any verbal instructions or verbal responses and thus, is useful in assessing the early childhood population. The test contains two subtest categories: a Cognitive Battery and an Attention/Memory Battery. The Cognitive Battery is comprised of five subtests of nonverbal intellectual ability related to visualization and reasoning with four subtests required to obtain a Nonverbal IQ composite score. The Attention/Memory Battery also has five subscales that combine to create a composite score for Nonverbal Memory and Processing Speed. Authors of the Leiter-3 provide substantial evidence for content, criterion, concurrent, and construct validity, specifically citing studies (Grondhuis & Mulick, 2013) demonstrating convergent validity of the Leiter-3 with another commonly used intelligence test, the Stanford Binet-5 (SB-5), in ASD populations (Buros Center for Testing, 2014; Roid et al., 2013).

Autism Diagnostic Observation Schedule, Second Edition (ADOS-2). The ADOS-2 is a semi-structured, standardized assessment administered by a trained examiner that aims to assess an individual's use of social, play, and communication (Lord, Rutter, et al., 2012). The examiner conducting the assessment implements a series of standardized "presses" that are purposed to elicit communicative behaviors from the examinee. There are five modules in total, and module selection is based on the individual's level of expressive language and age. Module 1 is utilized for individuals with no language or only single words; Module 2 is utilized for individuals who demonstrate phrase speech but are not yet verbally fluent; Module 3 is utilized for children and adolescents who are verbally fluent; Module 4 is utilized for older adolescents and adults who are verbally fluent; and the Toddler Module is utilized for children between 12-30 months of age. The measure is comprised of two domains, Social Affect (SA) and Restrictive, Repetitive Behaviors (RRB), that combine to produce a Total Score. This total score can then be used to compare the severity of autism symptoms across Module types and is often reported as the primary measurement value in autism research (Gotham, Risi, Pickles, & Lord, 2007; Hus, Gotham, & Lord, 2014; Le Couteur, Haden, Hammal, & McConachie, 2008; Lord, Rutter, et al., 2012).

In the context of this study, the Toddler Module, Module 1, and Module 2 were most frequently administered. We used the algorithm total score (a composite of the two subdomain scores, Social Affect and Restrictive, Repetitive Behaviors) to examine the relationships between this total score and the scores obtained on the BASC-3 PRS-P to

investigate concurrent validity of the two outcome values. The Total Score value on the ADOS-2 is thought to be the best overall estimate of autism symptomology and is most used across studies of autism research (Gotham et al., 2007; Hus et al., 2014; Lord et al., 2014).

Psychometrically, the ADOS-2 demonstrates moderate to high reliability and validity, ranging across module types (Hus et al., 2014). The Social Affect (SA) is reported to have high internal consistency (Cronbach's $\alpha > .85$ across Modules 1- 3) while the Restricted, repetitive behavior (RRB) domain shows poor internal consistency (Cronbach's $\alpha > .47$ across Modules 1- 3). For the Toddler Module, the authors report a range of poor to high internal consistency (Cronbach's $\alpha > .50-.90$ across SA and RRB domains) (Lord, Luyster, Gotham, & Guthrie, 2012; McCrimmon & Rostad, 2013). Similarly, test-retest reliability has been reported to be adequate to high across domains and total scores (.64 -.92) across Toddler Module- Module 3. Agreement in diagnostic classification is high and ranges from 92 – 98%. Items were selected for the algorithm based upon exploratory and confirmatory factor analyses resulting in the finding that items were adequately correlated with each other, $< .70$. Sensitivity and specificity for differentiating ASD are reported to be between 60%-95% and 75%-100%, respectively. The toddler module reported 87% sensitivity and between 86-91% specificity across diagnostic groups (McCrimmon & Rostad, 2013).

Data analyses

All analyses were performed using SPSS version 26.0.

CHAPTER FOUR

Results

Preliminary Findings

Prior to main analyses, all variables were examined through IBM SPSS program version 26.0 for accuracy of data entry, missing values, and fit between their distributions. There were several notable findings: There was a poor split between Spanish and English Language BASC-3 PRS-P questionnaires within the sample (10 Spanish to 105 English), as well as between males vs females within the total sample and within the diagnostic subgroups (97 males to 18 females in the total sample; see Table. 3 for group proportions). Additionally, data for two independent variables, Chronological Age and Gender, violated basic assumptions of normality. For both variables, appropriate log transformations were performed; however, results of these transformations did not yield improvements. As such, nonparametric statistics were used when conducting group comparisons and associations that included these variables.

Regarding the split between Spanish and English language cases of the BASC-3 PRS-P within the sample, it was determined that due to the small amount of Spanish cases, statistical analyses examining the reliability of the DSD scale in Spanish language questionnaires would yield unmeaningful (and potentially invalid) conclusions. Based on this reasoning, we did not examine the reliability for the DSD scale for the Spanish language cases any further. Internal consistency of the DSD scale on English language questionnaires within the sample was examined and reported below.

Regarding gender differences, it was hypothesized that although there was a small number of females overall ($n=18$), the proportion of males to females within the sample may still be representative to gender proportions observed in the general population of children with ASD, as it has been repeatedly established that ASD is observed higher in males than females (Halladay et al., 2015; Ros-Demarize et al., 2020). If gender proportions reported in the literature for 4-year-olds with ASD were not statistically different than the current sample's gender proportions, this would provide increased confidence in interpreting and extrapolating potential findings related to gender differences within the study. The current national prevalence rates of 4-year-old children with ASD report between a 2.6 – 5.2:1 male-to-female ratio (Christensen et al., 2019), while the male-to-female ratio in the current study is 5.3:1. A chi-square goodness-of-fit test was conducted to test whether the ratio of males-to-females in the current sample was statistically different than the ratios reported in the literature. Results indicated there were no significant differences in the proportion of males to females identified in the current sample (5.3:1) as compared with the upper range ratio of (5.2:1) that was obtained in a previous nationwide study, $\chi^2(1, n = 115) = .838, p = .36$. Based on these findings, we continued with further analyses examining gender differences on the DSD scale.

Descriptive Statistics. Overall sample demographic characteristics are reported below in Table 3 in terms of diagnostic group. The study consisted of 97 male and 18 female preschool-aged children, who were predominantly White/Caucasian, Non-Hispanic. The average age at time of evaluation and completion of the BASC-3 PRS-P

questionnaire was 45 months (range = 24 - 70 months, SD = 13 months). There were 67 children (58.3%) in the younger preschool age-group, which consisted of children ages 2-3, while 47 children (40.9%) were in the older age group, consisting of children ages 4-5. ASD and Non-ADS groups were similar in terms of their mean chronological age, age range, and gender proportions.

Table 3.
Descriptive Statistics for Demographic Variables by Groups

	ASD (N=75)	Non-ASD (N=40)	Total
Age (Months)			
Mean (SD)	45.53 (13.28)	45.50 (12.79)	45.52 (13.06)
Age Range (Months)	24–68	26–70	24-70
	<u>N (%)</u>	<u>N (%)</u>	
Sex/Gender			
Female	12 (16 %)	6 (15%)	18 (15.7%)
Male	63 (84 %)	34 (85 %)	97 (84.3%)
Race			
Caucasian	58 (77.3 %)	29 (72.5 %)	87 (75.7%)
African American	8 (10.7%)	4 (.01 %)	12 (10.4%)
Asian	4 (5.3 %)	1 (2.5 %)	5 (4.3%)
Native Hawaiian/Pacific Islander	-- --	1 (2.5 %)	1 (0.9%)
Multiracial	4 (5.3 %)	2 (5%)	6 (5.2%)
Unknown/Not Reported	1 (1.3%)	3 (7.5 %)	4 (3.5%)
Ethnicity			
Hispanic	27 (36%)	15 (37.5%)	42 (36.5%)
Non-Hispanic	48 (64%)	23 (57.5%)	71 (61.7%)
Unknown	-- --	2 (5%)	2 (1.7%)

Group Differences.

Group Differences on Demographic Characteristics. Several analyses were conducted to identify potential differences in subject characteristics amongst the two diagnostic groups (ASD vs. Non-ASD). Chi square analyses (with Yates' Continuity Correction) revealed no significant differences between group type and gender, $\chi^2 (1, n = 115) = .00, p = 1.0, \phi = .013$, as well as ethnicity, $\chi^2 (1, n = 115) = .024, p = .877, \phi = -.034$. With nine expected cell counts less than five, Fisher's exact test (2 x c) was conducted to examine whether the proportion of races were different between diagnostic groups. Results indicated a non-significant difference in proportions, $p = .42$. Chronological age was found to be non-normally distributed within the sample as a whole and natural log transformation was performed to potentially enhance the distribution; however, results did not yield improvements. As such, a nonparametric, Mann-Whitney U Test was conducted to determine if there were differences in chronological age between diagnostic groups. Results indicated no significant difference in the chronological age for ASD group ($Mdn = 44$) and the Non-ASD group ($Mdn = 43$), $U = 1473.5, p = .967$.

Group Differences on Study Measures. Mean (M) and standard deviation (SD) values for relevant variables on study measures are listed below in Table 4.

Table 4.
Group Characteristics on Study Measures

	ASD (<i>N</i> = 75)		Non-ASD (<i>N</i> = 40)	
	<i>N</i>	<i>M (SD)</i>	<i>N</i>	<i>M (SD)</i>
Combined Nonverbal DQ/IQ ^a	68	71.30 (23.20)	38	83.04 (23.24)
MSEL				
Visual Reception Age Equivalent (Months)	53	26.47 (11.17)	26	30.69 (14.50)
Fine Motor Age Equivalent (Months)	53	27.19 (10.29)	26	30.85 (12.75)
Nonverbal Developmental Quotient (NVDQ)	53	64.81 (20.13)	26	75.45(22.85)
DAS-II				
Nonverbal Reasoning Composite	13	92.23 (19.37)	12	99.5 (14.06)
ADOS-2				
Total Score	70	18.29 (5.89)	39	9.79 (6.43)
SA Score	65	12.08 (5.20)	38	7.42 (5.59)
RRB Score	65	5.31 (2.18)	38	3.66 (2.28)
BASC-3 DSD Scale				
General Norms	75	72.21 (12.58)	40	75.65 (12.95)
Clinical Norms	74	57.73 (9.15)	40	58.00 (8.3)
Gender Specific, Clinical Norms	74	57.24 (9.06)	40	57.68 (8.27)

MSEL: Mullen Scales of Early Learning, Nonverbal Developmental Quotient presented as Standard Scores.

DAS-II: Differential Ability Scales, Second Edition, presented as Standard Scores.

^a*Combined Nonverbal DQ/IQ*: Estimated Nonverbal Developmental Quotient or Nonverbal IQ across MSEL, DAS-II, and Leiter-3 scores, presented as Standard Scores.

DSD: Developmental Social Disorders. BASC-3 DSD Scale scores presented as T-scores.

ADOS-2: Autism Diagnostic Observation Schedule, Second Edition.

ADOS-2 SA: ADOS-2 Social Affect score

ADOS-2 RRB: ADOS-2, Restricted, Repetitive Behaviors score

Independent samples t-tests were conducted to determine whether there were mean differences between groups on variables among study measures. Results from these

revealed no significant differences between diagnostic groups across the following variables: age equivalent scores on the Visual Reception (VRae) and Fine Motor (FMae) for the Mullen Scales of Early Learning, $t(77) = 1.423, p = .158$; $t(77) = 1.37, p = .175$; and the DAS-II nonverbal reasoning scores, $t(23) = 1.07, p = .298$. The magnitude of the differences in the means for VRae ($M = 4.22$, 95% CI [-1.67 to 10.11]) was small (Cohen' $d = 0.325$); FMae ($M = 3.65$, 95% CI [-1.65 to 8.97]) was small (Cohen' $d = 0.315$); and DAS-II Nonverbal Reasoning ($M = 7.26$, 95% CI [-6.83 to 21.37]) was also small (Cohen' $d = 0.43$).

In contrast, significant differences were found between the ASD and Non-ASD groups in terms of nonverbal developmental/cognitive level of functioning, with the ASD group demonstrating significantly lower nonverbal cognitive abilities, $t(76.56) = 2.50, p = .015$. The magnitude of the differences in the means for Nonverbal Developmental Quotient/Intelligence Quotient (NVDQ/IQ) between groups, $M = 11.74$, 95% CI [-2.37 to 21.11], was medium (Cohen' $d = 0.51$). Similarly, significant mean differences were found between diagnostic groups for all the composite scores of the ADOS-2. As expected, the ASD group scored significantly higher than the Non-ASD group the Social Affect (SA) score, $t(101) = -4.26, p < .001$, the Restricted, Repetitive Behavior (RRB) score, $t(101) = -3.64, p < .001$, and the ADOS-2 Total Score, $t(107) = -7.05, p < .001$.

BASC-3 PRS-P DSD Scale

The BASC-3 PRS-P DSD scale was examined across several different psychometric areas to gain a more complete understanding of the scale's performance within this sample and to test its discriminative accuracy. We investigated aspects of the DSD scale's reliability, examined group differences on the scale, reported correlations with other relevant variables (i.e., ADOS-2 scores), and examined the scale's discriminative ability to classify ASD vs. Non-ASD cases. Upon initial review and investigation of the sample's DSD T-scores (as calculated according to the BASC-3 PRS-P's normative scores from similar-aged children across the U.S. with both typical and atypical development), we subsequently considered the potential usefulness of further examining the DSD scale's classification abilities when calculating DSD scores using two additional normative groups, a clinical norm group comprised of children ages 4-5 with atypical development (i.e., diagnosed with ADHD, Language Delay, ASD, conduct problems, etc.), and a clinical normative group separated by gender. We believed that examining the sample's DSD scores in comparison to scores obtained by each normative group would ultimately yield more comprehensive understanding of the utility of the DSD within a clinically referred population. Results from these analyses are detailed below.

Reliability. We initially proposed to examine the internal consistency of the DSD scale for both Spanish and English language versions of the BASC-3 PRS-P within the sample. Due to the small number cases of Spanish language versions, Cronbach's alpha

comparisons between Spanish and English versions unable to be determined. According to Reynolds and Kamphaus (2015), the Developmental Social Disorder scale (English Language version) has good internal consistency, with a Cronbach alpha coefficient of .85. Preferably, the Cronbach alpha coefficient of a scale should be above .7 (DeVellis, 2016). In the current study, the Cronbach alpha coefficient was .84 for English language versions of the BASC-3 PRS-P DSD scale, indicating that the 17 items included in the DSD scale have a high level of internal consistency and are reliably measuring the same underlying construct associated with developmental social disorders.

DSD Scale Group Differences. Mean differences were examined between the ASD and Non-ASD group based on the obtained T-scores when using the following BASC-3 PRS-P scoring comparison groups: the general normative sample, the clinical normative scores, and the gender specific clinical normative sample. Results of these comparisons are listed below in Table 5. Overall, mean DSD scores did not significantly differ between ASD and Non-ASD groups when scores were calculated using each of the different normative comparison groups.

Table 5.*T-Tests for BASC-3 DSD Scale Utilizing Different Normative Comparison Groups*

BASC-3 Scale	ASD (n=74) Mean (SD)	Non- ASD (n=40) Mean (SD)	<i>F</i> - value	<i>p</i> -value	Cohen's <i>d</i>
DSD Scale General Norms	72.21 (12.58)	75.65 (12.95)	1.38	.170	0.27
DSD Scale Clinical Norms	57.73 (9.15)	58.00 (8.29)	.15	.877	0.03
DSD Scale Gender Specific, Clinical Norms	57.24 (9.06)	57.68 (8.27)	.25	.803	0.05

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

DSD Scale Correlations. Additionally, we examined relationships between the scores obtained on the DSD scale and participant characteristics, as well as scores obtained on other study measures. The continuous variables ‘chronological age’ and ‘gender’ did not meet assumptions of normality in the total sample. As previously mentioned, log transformations were applied in attempts to better normalize the data but did not result in improvements. Due to this, nonparametric Spearman's rank-order correlation was conducted to examine the relationships between DSD scores and chronological age, nonparametric Kendall's Tau-b was used for correlational analyses for gender and ethnicity, while Pearson's product-moment correlation was used with the remaining demographic variables. Kendall's Tau-b was selected as the nonparametric test of choice over Spearman's Rho when correlational tests were to be conducted between a

categorical, dichotomous variable (such as gender and ethnicity) and continuous variables (Arndt, Turvey, & Andreasen, 1999; Marascuilo & McSweeney, 1977). Results from correlational analyses are detailed below and correlation coefficient values are summarized in Table 6.

Chronological Age. There were significant negative correlations between chronological age and DSD scale T-scores: DSD General Norm T-Scores, ($r_s = -.43$, $p < .001$), DSD Clinical Norm T-Scores ($r_s = -.28$, $p = .002$), and DSD Gender Specific Clinical Norm T-Scores, ($r_s = -.28$, $p = .002$). Overall, as age increased within the sample, DSD T-scores decreased.

Gender. Gender was not strongly associated with any of the DSD Norm T-scores: DSD General Norm T-Scores, ($\tau b = -.08$, $p = .276$); DSD Clinical Norm T-Scores, ($\tau b = .00$, $p = .991$); DSD Gender Specific Clinical Norm T-Scores, ($\tau b = .05$, $p = .519$). Associations between gender and DSD scores were both weak and statistically non-significant; however, when taken at face-value, males obtained higher mean DSD T-scores when scores were calculated using both the DSD General Norm comparison group (Males, $M(SD) = 73.95 (1.23)$; Females, $M(SD) = 70.50 (3.73)$), and the DSD Clinical Norm group (Males, $M(SD) = 57.90 (.88)$; Females, $M(SD) = 57.44 (2.43)$). Interestingly, when using the DSD Gender Specific Norm group to calculate scores, females obtained slightly higher DSD T-scores than males (Males, $M(SD) = 57.18 (.84)$; Females, $M(SD) = 58.56 (2.71)$).

Race and Ethnicity. Race and ethnicity also showed weak, non-significant associations with the DSD scores. The following results reflect analyses of race by DSD T-Scores: DSD General Norm T-Scores, ($r = -.02, p = .793$); DSD Clinical Norm T-Scores, ($r = -.11, p = .238$); DSD Gender Specific Clinical Norm T-Scores, ($r = -.10, p = .254$). Examining ethnicity by DSD T-scores resulted in: DSD General Norm T-Scores, ($r = .08, p = .392$); DSD Clinical Norm T-Scores, ($\tau b = .08, p = .283$); DSD Gender Specific Clinical Norm T-Scores, ($\tau b = .07, p = .330$). Nonparametric Kendall's Tau-b was used for correlational analyses when examining the relationship between ethnicity and DSD Clinical T-Scores and DSD Gender Specific Clinical T-Scores, as both distributions contained a significant degree of heterogeneity in their variances.

Nonverbal Developmental/Cognitive Level. There were also no significant correlations found between DSD scores and nonverbal developmental quotient/intelligence quotient (NVDQ/IQ): DSD General Norm T-Scores, ($r = -.05, p = .622$); DSD Clinical Norm T-Scores, ($r = -.03, p = .781$); DSD Gender Specific Clinical Norm T-Scores, ($r = -.02, p = .839$).

Table 6.

Correlation Matrix of Participant Characteristic Variables by BASC-3 DSD Scale T-Scores Using General, Clinical, and Gender Specific- Clinical Normative Scoring in total sample^a

Variable	DSD Scale General Norms	DSD Scale Clinical Norms	DSD Scale Gender Specific, Clinical Norms
1. Age	-.43 ** (114)	-.28** (114)	-.28** (114)
2. Gender	-.08 (115)	.00 (114)	.05 (114)
3. Race	-.02 (115)	-.11 (114)	-.10 (114)

4. Ethnicity	.08 (113)	.08 (112)	.07 (112)
5. NVDQ/IQ ^b	-.05 (106)	-.03 (106)	-.02 (106)

* $p < .05$, ** $p < .01$

^aNumbers in parentheses equal N.

^bNonverbal Developmental Quotient/Intelligence Quotient.

Concurrent Validity of the DSD Scale with ADOS-2 Scores

In order to assess for concurrent validity, Pearson product-moment correlations were conducted to identify whether there were significant associations between the scores on the BASC-3 Developmental Social Disorders (DSD) scale (T-scores calculated using General Norms, Clinical Norms, and Gender Specific Clinical Norms) and scores from the ADOS-2, including the Social Affect (SA) scores, Restricted, Repetitive Behaviors (RRB) scores, and Total scores. Preliminary analyses showed linear relationship and normal distributions among variables, as assessed by Shapiro-Wilk's test ($p > .05$), with no outliers. Overall, there were very weak correlations between DSD scores and ADOS-2 scores, indicating poor concurrent validity between the scales. Stronger associations between the DSD scale and ADOS-2 scores were expected given previous results of moderate, positive correlations ($r=.30$) between BASC-2 DSD scale scores and an ADOS and ADOS-2 severity score reported by Bradstreet et al. (2017). Associations between DSD scores and ADOS-2 Social Affect (SA) scores, Restricted, Repetitive Behaviors scores (RRB), and Total scores are outlined below, and a summary of correlation coefficient values are listed in Table 7.

Social Affect scores x DSD T-Scores. Associations between SA scores and DSD T-scores were as follows: DSD General Norm T-Scores, $r(103) = -.07, p = .486$; DSD Clinical Norm T-Scores, $r(103) = -.10, p = .335$; and DSD Gender Specific Clinical Norm T-Scores, $r(103) = -.08, p = .401$.

Restricted, Repetitive Behaviors scores x DSD T-Scores. Correlations between RRB scores and DSD T-scores were as follows: DSD General Norm T-Scores, $r(103) = -.07, p = .510$; DSD Clinical Norm T-Scores, $r(103) = -.08, p = .424$; and DSD Gender Specific Clinical Norm T-Scores, $r(103) = -.07, p = .476$.

Total Scores x DSD T-Scores. Correlations between ADOS-2 Total scores and DSD T-scores were as follows: DSD General Norm T-Scores, $r(109) = -.06, p = .524$; DSD Clinical Norm T-Scores, $r(109) = -.09, p = .329$; and DSD Gender Specific Clinical Norm T-Scores, $r(109) = -.09, p = .375$.

Table 7.

Pearson's Correlation Matrix of BASC-3 DSD Scale T-Scores (General, Clinical, and Gender Specific Clinical Normative Scoring) with ADOS-2 Scores in the total sample^a

Measure	ADOS-2 SA ^a Scores	ADOS-2 RRB ^b Scores	ADOS-2 Total Scores
DSD T-Scores General Norms	-.07 (103)	-.10 (103)	-.08 (109)
DSD T-Scores Clinical Norms	-.07 (103)	-.08 (103)	-.07 (109)
DSD T-Scores Gender Specific, Clinical Norms	-.06 (103)	-.09 (103)	-.09 (109)

^a $p < .05$, ^{**} $p < .01$.

^a Numbers in parentheses equal N.

^b Social Affect (SA); ^c Restricted, Repetitive Behaviors (RRB).

Discriminative Accuracy of the DSD Scale for Diagnostic Classification.

Logistic Regression Analyses. Several direct logistic regression analyses were performed to determine the influence of DSD T-Scores on the likelihood that participants were diagnosed with ASD or another developmental delay (Non-ASD). We conducted three sets of analyses on the groups of DSD T-Scores calculated using the General normative sample, the Clinical normative sample, and the Gender-Specific Clinical normative sample. For all distributions, linearity of the continuous variables (DSD scores) with respect to the logit of the dependent variable (diagnostic group type) was assessed via the Box-Tidwell (1962) procedure. Based on results of this procedure, all continuous independent variables were found to be linearly related to the logit of the dependent variable. No outliers were identified within the data.

DSD T-Scores Using the General Normative Comparison Group. An overall logistic regression model consisting of one continuous, predictor variable (DSD General T-Scores) and one dichotomous, outcome variable (diagnostic group) was created to assess the influence of DSD T-Scores (calculated using the general normative comparison group) on the probability for correct diagnostic classification of ASD vs. Non-ASD preschoolers. Three, separate logistic regressions were conducted at cutoff values of 60 (Juechter, 2012), 63 (Bradstreet et al., 2017) and 70 (Reynolds & Kamphaus, 2015), as these values were previously found to yield optimal sensitivity/specificity when

using the BASC-2 DSD scores to distinguishing between preschoolers with ASD from typically developing children or children with other developmental delays.

Results of the logistic regression model evaluating the ability of the DSD to accurately identify children with ASD and differentiate them from children with other diagnoses revealed non-significant results, $\chi^2(1) = 1.913, p = .167$. A cut score of 60 found a 65.2 % accuracy rate with 84% of participants identified as having an autism spectrum disorder (Sensitivity), and accurately screened out 30% of participants without an ASD (Specificity). A cut score of 63 accurately classified 71% of preschoolers with ASD while accurately screening out 48% of Non-ASD participants. A cut score of 70 produced an overall PAC hit-rate of only 44%, and accurately screened in 25% of participants with an ASD while 78% of participants without an ASD were accurately screened out.

DSD T-Scores Using the Clinical Normative Comparison Group. We repeated the previously outlined logistic regression procedures, examining the influence of the DSD T-Scores on the probability of ASD vs. Non-ASD classification; however, these analyses utilized DSD T-Scores that were calculated using a Clinical norm comparison sample. These analyses also revealed non-significant results, $\chi^2(1) = .025, p = .875$. A cut score of 60 found a 64.9 % accuracy rate with 100% of participants identified as having an autism spectrum disorder (Sensitivity), and accurately screened out 0% of participants without an ASD (Specificity). A cut score of 63 similarly classified 100% of preschoolers with ASD while failing to screen out all Non-ASD

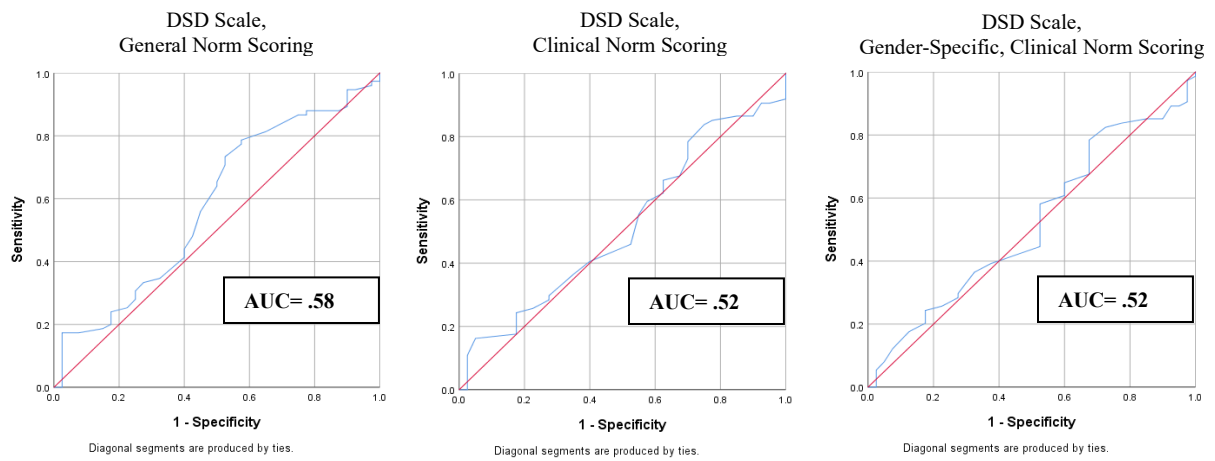
participants. Finally, a cut score of 70 produced an overall PAC hit-rate of only 35.1%, and accurately screened in 0% of participants with an ASD, while 100% of participants without an ASD were accurately screened out.

DSD T-Scores Using the Gender-Specific Clinical Normative Comparison Group. This set of DSD T-Scores were calculated using the gender-specific, clinical normative comparison group. A cut score of 60 found a 64.9 % accuracy rate with 100% of participants identified as having an autism spectrum disorder (Sensitivity) and failed to screen out all participants with Non-ASD diagnoses (Specificity). A cut score of 63 produced the same sensitivity and specificity values as observed using a cut score of 60. Finally, a cut score of 70 produced the lowest overall PAC hit-rate of only 31.5%, and accurately screened in 0% of participants with an ASD, while 100% of participants without an ASD were accurately screened out.

Receiver Operating Characteristic (ROC) Curves. Receiver Operating Characteristic (ROC) analysis is used to examine the performance of a diagnostic test as well as to evaluate the diagnostic accuracy of other statistical models (Zou, O'Malley, & Mauri, 2007), such as logistic regression analyses. Specifically, the ROC Curve is a visual representation a plot of sensitivity values against (1-specificity) values, and produces an Area Under the Curve (AUC) measure “that averages diagnostic accuracy across the spectrum of test values” (Zou et al., 2007). The AUC value can range from 0.5 to 1.0 and higher values indicate better discrimination of the statistical model (Laerd Statistics, 2017). To illustrate the findings from the logistic regression analyses, ROC

curves and AUC values were generated based on logistic regression results for each set of DSD T-Scores (General Normative Group, the Clinical Normative Group, and Gender-Specific, Clinical Normative Group). Figure 1 depicts the ROC Curves for logistic regression analyses corresponding to each DSD T-Score group.

Figure 1. ROC Curves: ASD vs. Non-ASD



According to Hosmer Jr, Lemeshow, and Sturdivant (2013), AUC values ranging in between 0.5-0.7 indicates poor discrimination, AUC values 0.7-0.8 indicates “acceptable” discrimination, 0.8-0.9 indicates “excellent” discrimination, and 0.9> indicates outstanding discrimination. For the DSD T-Scores calculated from the General norm group, the area under the ROC curve was .58, 95% CI [.468, .692], indicating poor discrimination. Poor discrimination was also found for the DSD T-Scores calculated from the Clinical norm group, AUC= .517, 95% CI [.406, .627], and the Gender-Specific, Clinical norm group AUC=.517, 95% CI [.407, .629]. These results are in line with the non-significant findings from the logistic regression models performed at given cut-

points and provide an additional indicator that the DSD scales are a poor measure of diagnostic discrimination within this sample.

Sensitivity, Specificity, and Likelihood Ratios. In addition to conducting logistic regression analyses and ROC curves as ways to examine diagnostic validity of the DSD scale, sensitivity and specificity values, and likelihood ratios (LR+/LR-) were also calculated to better examine trade-offs among a larger range of cutoff scores. Cutoff scores ranging from 60 to 70 were explored, as the BASC-3 manual reports DSD T-scores between 60-69 are considered in the “at risk” range and T-scores of 70 or higher are considered “clinically significant” (Reynolds & Kamphaus, 2015). Tables 8-10 below summarizes sensitivity/specificity values and likelihood ratios for each set of DSD T-Scores (General, Clinical, and Gender-Specific Clinical) in order to compare trade-offs in discriminative abilities and diagnostic accuracy.

Table 8.

Psychometrics of multiple cut-scores on the BASC-3 DSD content scale, T-Scores calculated using the General normative comparison group. ASD= 75, Non-ASD= 40.

DSD Cut-Score (T-Score)	Sensitivity, %	Specificity, %	LR+	LR-
60	84.0	30.0	1.20	0.53
61	81.3	35.0	1.25	0.53
62	77.3	42.5	1.35	0.53
63	70.7	47.5	1.35	0.62
64	65.3	50.0	1.31	0.69
65	60.0	52.5	1.26	0.76
66	48.0	57.5	1.13	0.90

67	44.0	60.0	1.10	0.93
68	34.7	67.5	1.07	0.97
69	30.7	75.0	1.23	0.92
70	25.3	77.5	1.13	0.96

Table 9.

*Psychometrics of multiple cut-scores on the BASC-3 DSD content scale,
T-Scores calculated using the Clinical normative comparison group.
ASD= 74, Non-ASD= 40.*

DSD Cut-Score (T-Score)	Sensitivity, %	Specificity, %	LR+	LR-
60	100.00	0.00	--	--
61	100.00	0.00	--	--
62	100.00	0.00	--	--
63	98.60	0.00	0.99	--
64	86.50	10.00	0.96	1.35
65	45.90	47.50	0.88	1.14
66	5.40	97.50	2.16	0.97
67	0.00	100.00	--	--
68	0.00	100.00	--	--
69	0.00	100.00	--	--
70	0.00	100.00	--	--

Table 10.

*Psychometrics of multiple cut-scores on the BASC-3 DSD content scale,
T-Scores calculated using the Gender-Specific, Clinical normative comparison group.
ASD= 74, Non-ASD= 40.*

DSD Cut-Score (T-Score)	Sensitivity, %	Specificity, %	LR+	LR-
60	100.00	0.00	--	--
61	100.00	0.00	--	--
62	98.60	0.00	0.99	--
63	93.20	0.25	0.96	2.70
64	82.40	27.50	1.14	0.64
65	44.60	0.47.50	0.85	1.17
66	21.60	82.50	1.24	0.95
67	0.00	97.50	0.00	1.03
68	0.00	100.00	--	--
69	0.00	100.00	--	--
70	0.00	100.00	--	--

CHAPTER FIVE

Discussion

The purpose of the present study was to determine the clinical utility of the DSD scale on the BASC-3 Parent Rating Scales, Preschool form to differentiate between children diagnosed with ASD from those with other developmental delays within a clinical sample of preschoolers. As the prevalence of ASD continues to rise, early identification practices must be improved to allow children access to early interventions (Maenner et al.; 2020). As many referrals for ASD diagnostic evaluations include children who have already been identified in medical or educational settings as having some form of developmental delay (Gordon-Lipkin et al., 2016), it is also increasingly important to identify diagnostic instruments that are efficient and have adequate specificity in differentiating ASD from the wide array of developmental disorders that can present in the early childhood period.

Correct identification is crucial as a diagnosis often informs which evidence-based interventions and clinical recommendations will be most effective for a given child's behavioral presentation. Many young children who present for diagnostic evaluations for ASD display atypical behaviors that overlap and are similar to other developmental diagnoses. One assessment method suggested to aid in screening and diagnosis of ASD in young children is the utilization of parent-informed, broadband behavioral questionnaires that examine an range of social, emotional, and behavioral

issues in early childhood (C. L. Myers et al., 2010; Rescorla et al., 2017; Reynolds & Kamphaus, 2015).

The current study investigated the use of the Behavior Assessment System for Children, Third Edition (BASC-3), to discriminate ASD symptoms in a clinical sample. To our knowledge, the use of the BASC-3 PRS-Preschool version has yet to be examined in a clinical population and has not been validated in ASD populations. We examined the clinical validity of the Developmental Social Disorders (DSD) scale on the BASC-3 PRS-P to meaningfully differentiate between preschoolers diagnosed with ASD and those diagnosed with other developmental delays. We examined and compared three sets of DSD scale T-Scores calculated based on distinct, normative comparison groups: a General Normative group, a Clinical Normative group, and a Gender-Specific Clinical Normative group. Evaluation of DSD score results across the three normative groups allowed for a more comprehensive investigation of the functionality of the DSD scale in a fully clinical population of preschool-aged children. Additionally, the study investigated aspects of reliability of the DSD scale within the sample, as well as the DSD scale's associations with other variables such as nonverbal level of cognitive functioning, age, gender, and total scores obtained on the ADOS-2, a commonly-used diagnostic instrument for ASD.

Developmental Social Disorders (DSD) Scale

Reliability. An initial aim of the study was to investigate aspects of reliability of the DSD scale within the present sample. The authors of the BASC-3, Reynolds and Kamphaus (2015), described the DSD scale as a scale that measures “the tendency to display behaviors characterized by deficits in social skills communication, interests, and activities; such behaviors may include self-stimulation, withdrawal, and inappropriate socialization.” In the current study, findings on the reliability of the DSD scale within the sample indicated a high level of internal consistency, with a Cronbach alpha coefficient of .84 for the English language version. These results suggest that the items from the scale are consistently attempting to measure the same underlying construct described above. Additionally, this finding was consistent with level of internal consistency reported in the BASC-3 manual for the Parent Rating Scale, Preschool Version, which was found to have an alpha value of .85.

Concurrent Validity with the ADOS-2. We examined the concurrent validity of scores obtained on the BASC-3 PRS-P, DSD scale with several total scores from the ADOS-2. As expected, strong, positive correlations were found between diagnostic group type and ADOS-2 SA, RRB, and overall total scores; with higher ADOS-2 scores associated with increased frequency of participants belonging to the ASD group. Contrary to our hypothesis that DSD T-scores would show small to moderate positive associations with ADOS-2 scores, all three sets of DSD scores (calculated using General, Clinical, and Gender-Specific Clinical comparison groups) showed non-significant, weak

associations with each ADOS-2 total score. Moreover, r values were *negative*, suggesting a decrease in DSD T-scores as ADOS-2 total scores became larger. This finding was surprising given that increased ADOS-2 scores are associated with a greater likelihood of having ASD. In line with this, DSD scores also demonstrated negative, weak associations with diagnostic group type.

In contrast to our findings, Bradstreet et al. (2017); Juechter (2012) reported small ($r=.254$), but significant, concurrent validity of the BASC-2 DSD scale and the ADOS/ADOS-2 total score. There may be several explanations for this discrepancy. First, the present study examined the updated DSD scale in a sample that was comprised entirely of clinically-referred preschool-aged children, while previous research assessed the scale in subgroups of preschoolers with ASD, other developmental delays, as well as in typically developing children. Though not impossible, it is highly unlikely that a typically developing child would be referred to a tertiary clinic for children with developmental disabilities, as there are multiple screening elements performed during the referral process. Thus, our sample characteristics may explain the discrepancy.

Secondly, there are differences in the type of ADOS scores utilized across studies. For instance, Juechter (2012) looked at subdomain scores from an older version of the ADOS (the ADOS-G), while the current study examined subdomain total scores obtained on the current version, the ADOS-2. Concurrent validity findings from Bradstreet et al. (2017) examined correlations between the DSD scale and a Calibrated Severity Score (CSS) on the ADOS-2, a different score than that used in the present

study. Potentially, the finding here highlights differences in scoring practices related to underlying construct development of previous instruments or different scores on the ADOS-2.

Discriminative Accuracy: Group Comparisons. The study also examined whether the BASC-3 PRS-P DSD scale could reliably and meaningfully differentiate between preschool-aged children with ASD and those with other developmental delays within a clinically referred sample. Overall, both the ASD and Non-ASD groups obtained mean DSD T-Scores (ASD, $M = 72.21$; Non-ASD, $M = 75.65$) that were considered to be in the “clinically significant” range, according to qualitative score classifications outlined in the BASC-3 manual (Reynolds & Kamphaus, 2015); however, when comparing the groups, preschoolers in the ASD group did not obtain significantly different DSD scores than those in the Non-ASD group when scores were calculated using the General Normative group as the comparison sample (Figure 2). Similarly, there were no differences found between groups when DSD T-Scores were calculated using the Clinical or Gender-Specific Clinical Normative comparison groups (Figures 3 and 4).

Figure 2. Scatterplot of DSD Scores by Group Using General Norms

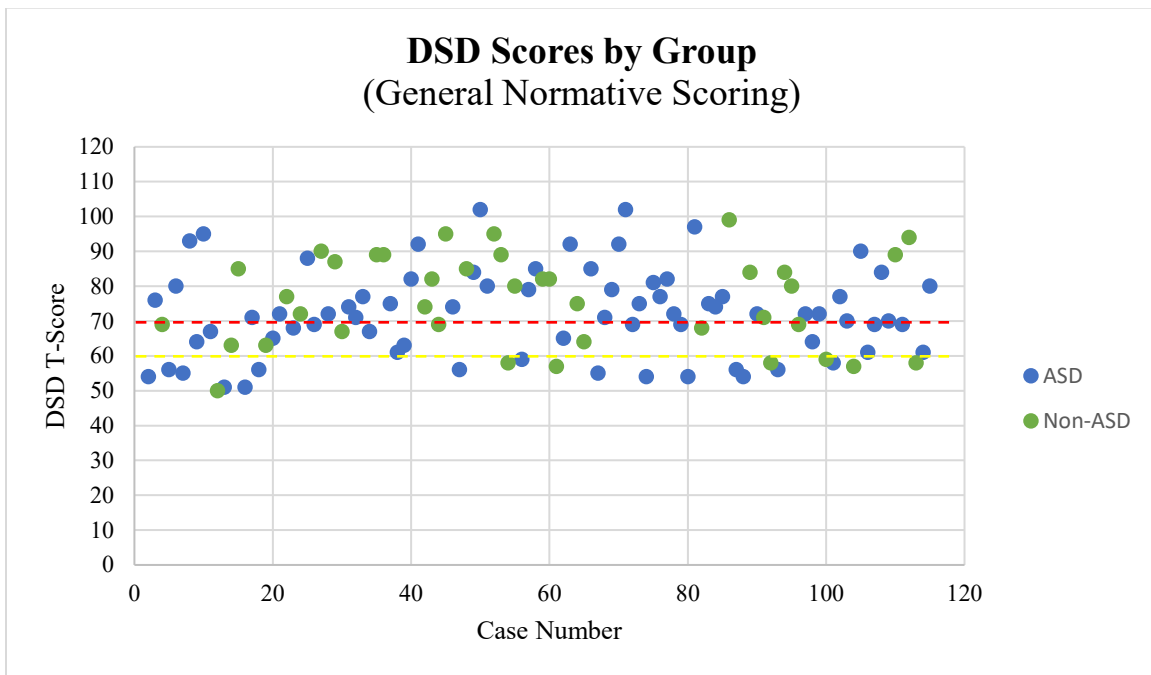


Figure 3. Scatterplot of DSD Scores by Group Using Clinical Norms

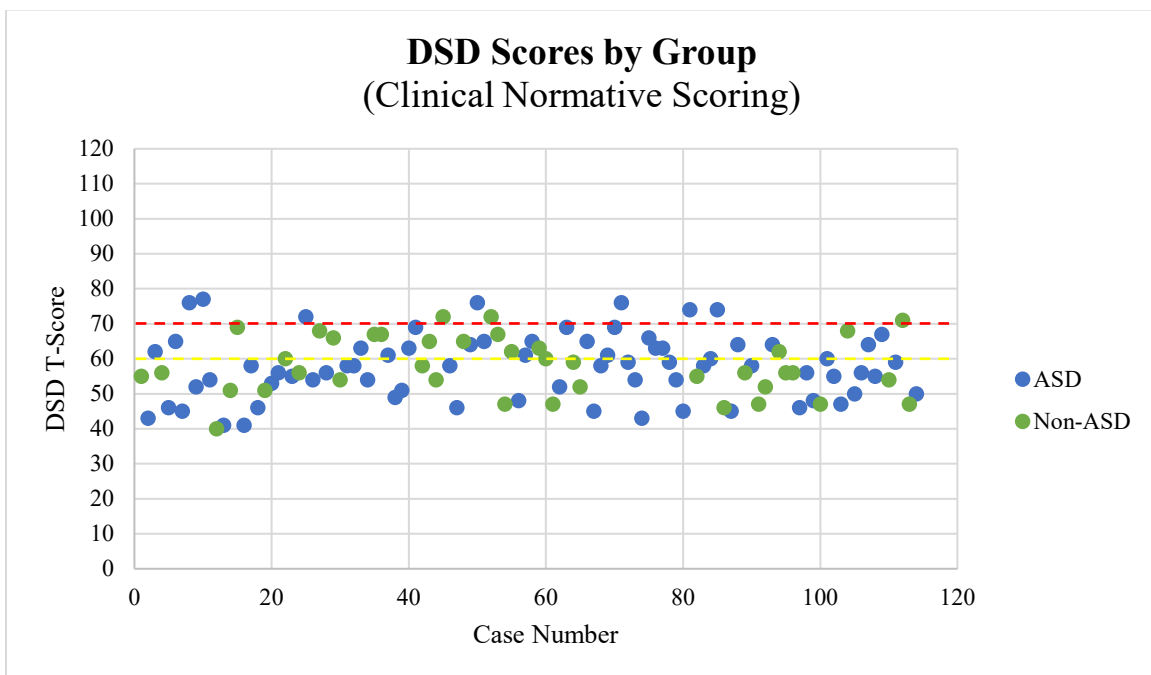
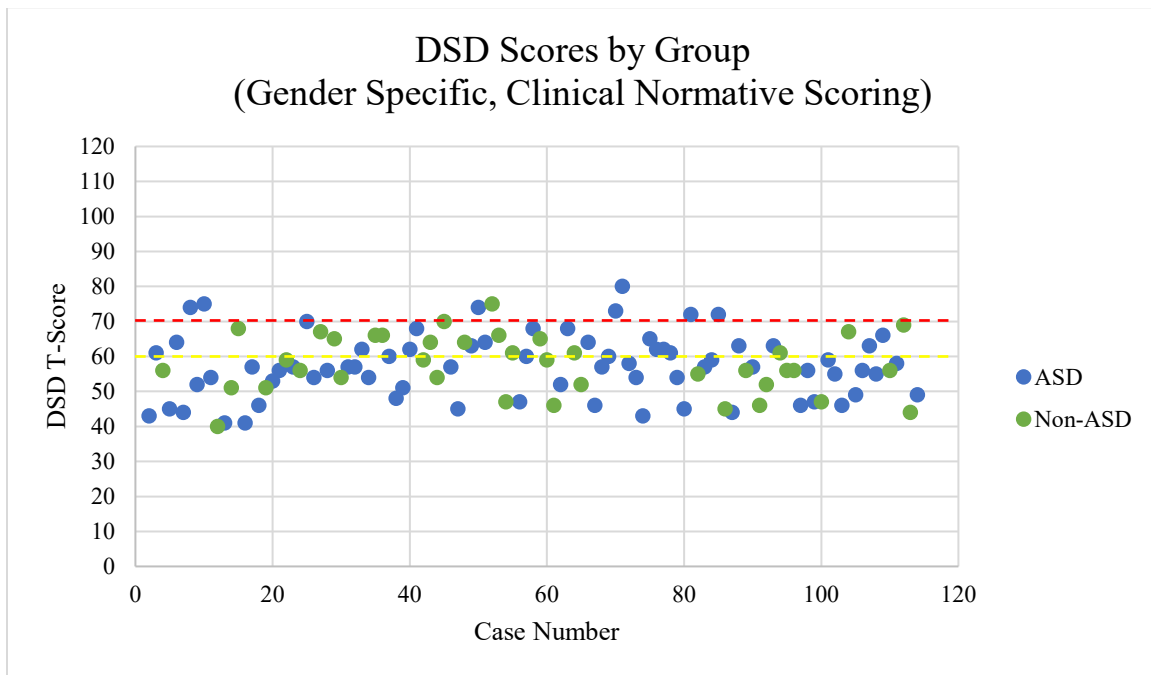


Figure 4. Scatterplot of DSD Scores by Group Using Gender-Specific, Clinical Norms



These results are consistent with previous research conducted by Gardner et al. (2017) which similarly found no differences in DSD scores between ASD and non-ASD groups (preschoolers diagnosed with an ID) on the BASC-2. However, other studies, including Juechter (2012) and Bradstreet et al. (2017), found ASD groups obtained significantly higher DSD scores than the non-ASD groups; however, it should be noted that results were significantly influenced by the composition of the participants in the ASD group, as in Juechter (2012), and in the comparison group in Bradstreet et al. (2017). Despite variations among statistical results, these studies, along with the current findings, document the difficulty in *reliably* differentiating ASDs from other developmental delays in early childhood due to the overlapping nature of behavioral symptoms.

Discriminative Accuracy: Group Prediction and Classification. The DSD scale showed low discriminative validity when used to differentiate preschoolers with ASD from those with other developmental delays in the present sample. Moreover, the scale was only slightly better than chance at correctly predicting diagnosis type. This finding was true for all sets of DSD T-Scores, regardless of which normative comparison sample was used for scoring; however, use of the General Normative sample produced the best discrimination, albeit poor overall. Consistent with past research investigating the use of broadband behavioral measures for screening and diagnostic purposes, the BASC-3 PRS-P DSD scale demonstrated that it is helpful in confirming the presence of developmental problems in participants but was unable to reliably distinguish between ASD and non-ASD membership.

For all screening and diagnostic measures, there is a clinical trade-off between sensitivity and specificity. The decision to prioritize one over the other is dependent upon the overall purpose of the measure. The current study chose to select the optimal cut score based primarily on a balance of sensitivity and specificity values, but slightly prioritizing sensitivity (true positives). This clinical decision would optimize rates of correctly identifying children who truly have ASD, but risks consequences of providing a portion of children with an incorrect ASD diagnosis. Despite this risk, the act of prioritizing sensitivity in selecting a cut off score for the DSD scale serves to minimize missing children who have a true ASD, while still allowing children who receive an incorrect ASD diagnosis the access to early intervention services, such as Applied

Behavior Analysis-based therapy, which has been shown to be beneficial to both children with ASD and other developmental delays (Feeley & Jones, 2006; Peters-Scheffer, Didden, Mulders, & Korzilius, 2010; Spreckley & Boyd, 2009).

An ideal screening or diagnostic instrument would have high levels of both sensitivity (.90 and greater) and specificity (also .90 or greater), meaning that the tool could accurately predict and detect both the true presence of a disorder as well as its true absence. While the initial ROC analyses revealed the DSD scale to be, overall, a poor measure at predicting ASD vs. Non-ASD group membership in the present sample of preschoolers, we proceeded to examine a range of cut scores and their corresponding sensitivity and specificity values to compare our findings to results from previous, BASC-2 DSD scale studies. When a cut score of 64 on the DSD scale was used (generated using the General Normative scoring), the scale accurately classified 65.3% of children with ASD (i.e., sensitivity= .65) and 50% of children without ASD (i.e., specificity = .50), values lower than those reported by Juechter (2012) and somewhat similar to values reported by Bradstreet et al. (2017). However, in efforts to best identify the participants with ASD within this sample, a better cut score selection would be a score of 60, thereby accurately classifying 84% of preschoolers with ASD, but only screening out 30% of preschoolers with other developmental delays.

Gender Differences. One final question the study aimed to answer was whether the optimal cut scores (60, 64) were equally valid at discriminating ASD and Non-ASD for both the boys and girls within the sample. We attempted to examine this by first

looking at correlations between gender and DSD scores, which demonstrated a weak relationship with each other. Then we further investigated whether gender significantly impacted the diagnostic prediction models when using a cut score of 60 and 64 and found that the gender of participants had little influence on increasing the accuracy of predicting correct ASD vs. non-ASD group membership. This was also true when prediction models were tested using DSD scores from the Clinical Normative and Clinical Gender Specific Normative scoring. Although the DSD scores obtained between males and females in the sample were not found to be statistically different, one notable finding was that females obtained slightly higher DSD T-scores than males (Males, $M(SD)= 57.18 (.838)$; Females, $M(SD)= 58.56 (2.705)$ when using the Clinical Gender Specific comparison group, possibly suggesting that the scale is more sensitive for detecting nuances in symptoms of atypical development in females when using this normative comparison group.

Clinical Implications for Practitioners

Findings from the current study have yielded several considerations for using the BASC-3 DSD scale as a tool to aid in differential diagnosis of preschoolers exhibiting signs of developmental delays. Initially, while further research is needed to verify and replicate these results, findings suggest that the DSD scale demonstrates adequate reliability within this clinical sample, indicating that the items included in the scale consistently measured what they were intended to measure (i.e., impaired social interactions, delayed development, atypical behaviors). However, the scale was not able

to meaningfully distinguish among different developmental disorders in preschool-aged children. Clinicians should be aware of the potential for low specificity when considering the DSD scale for use in differential diagnosis. Further, calculating DSD total scores based on the Clinical Normative and Clinical Gender Specific Normative comparison groups did not improve accurate distinction between preschoolers with ASD and those with other developmental delays – a surprising result given that the study was conducted within a fully clinical sample.

The BASC-3 PRS-P continues to show potential as a promising tool in the evaluation of preschool-aged children with developmental delays as it reliably detects symptoms of social and developmental impairment and promotes efficiency in the diagnostic process, ultimately facilitating early identification and initiation of intervention services for children with developmental concerns. Despite these abilities, the use of the DSD scale alone was shown to meaningfully differentiate between children with ASD and other developmental delays, thus cannot be used as a tool for earlier differential diagnosis within a clinical population. Due to the limited ability of the DSD scale to reliably discriminate between ASD and other early childhood disorders, practitioners may consider using results from the DSD scale to confirm the presence of atypical development and, perhaps, inform and guide the larger differential diagnostic process rather than using DSD scale results alone in determining diagnosis.

Strengths, Limitations, and Future Directions for Research

This study adds to the existing literature on use of behavioral rating scales in screening and diagnostic practices for identifying ASD in preschool-aged children. Another unique aspect of the project is that, to our current knowledge, no other studies have been conducted examining the updated version of the BASC-3 PRS-P in a clinical sample of preschoolers with developmental delays. In addition to these strengths, stringent criteria were required for clinical diagnosis of participants in each diagnostic group. Diagnoses of participants in the ASD group and in the other DD group were verified through Electronic Medical Records (EMRs) as having a comprehensive diagnostic evaluation that included the use of gold-standard assessment tools to help determine appropriate diagnosis. Moreover, diagnostic evaluations were conducted by highly trained clinical psychologists who specialize in the assessment of ASD in young children. Finally, another strength of the study is its attempt to address questions related to underrepresented subgroups within research samples, such as examining the Spanish Language version of the BASC-3 PRS-P and investigating gender differences in preschoolers with ASD and other developmental delays. Although there were small sample sizes of these subgroups (which resulted in difficulties producing valid statistical results), investigating the performance of screening and diagnostic measures in samples of minority groups is important in order to continue to promote equitable early identification practices and improve overall detection of ASD for all children.

In terms of limitations of the study, there are several notable issues. First, there was an extremely small number of Spanish-language BASC-3 forms that were completed, negating the possibility of conducting reliable statistical. Likewise, there were also a limited number of girls within the sample. Taken together, the generalizability of the statistical findings to the larger preschool-aged population of children with ASDs and other developmental delays is limited. While the number of girls in each group compared to boys was relatively representative of the larger gender prevalence ratio seen within neurodevelopmental disorders, the small number of females limits the true statistical confidence of the gender findings. Lastly, the present study was comprised of a smaller total sample size than previous studies investigating the BASC-2 in similar preschool populations. Replicating study methods in a larger sample would enhance confidence in the findings.

For future research, further comparisons of group scores on clinical and adaptive scales would be beneficial to determine whether there may be another scale on the BASC-3 PRS-P that better differentiates preschoolers with ASD from other developmental delays. For example, Reynolds and Kamphaus (2015) suggested that ASD may be a more accurate diagnosis when a) the DSD scale is elevated and the Conduct Problems and Aggression clinical scales are not, or b) when the DSD scale is elevated alongside the Withdrawal, Atypicality, and Attention Problems clinical scales. Additionally, although the present study looked at the concurrent validity of the DSD scale with the ADOS-2, a “gold-standard” diagnostic tool, there is a need for further

research examining the concurrent validity of the DSD scale specifically with other *parent-reported* diagnostic measures for ASD, such as the Social Communication Questionnaire (SCQ) or the Childhood Autism Rating Questionnaire (CARS-2). Examining other ASD-specific scales that employ the same method of assessment (parent-report) may lead to greater convergence than was shown with the ADOS-2. Lastly, it may be useful to compare items on the DSD scale and items on other ASD-specific, parent-report measures in order to obtain a better understanding of which DSD scale items may be more important for determining ASD versus developmental delay. This could potentially help inform the development of future BASC PRS preschool forms as the current BASC-3 preschool edition does not contain a specific ASD Probability scale, although one is included in the older child BASC PRS versions (BASC PRS, Children version and BASC PRS, Adolescent version) (Reynolds & Kamphaus, 2015). With these considerations in mind, researchers should continue to examine tools that would allow practitioners to accurately and efficiently diagnose ASD and pave the way towards early intervention.

Appendix A.

Table 1.

Measures of Behavioral and Psychiatric Comorbidities

Name	Abbreviation	Age Range	Description	Strengths/Weaknesses
Child Behavior Checklist 1.5-5 years	CBCL 1.5-5 years	18-72 months	Broadband, parent/caregiver rating scale comprised of 100 items that assess aspects of emotional and behavioral functioning such as emotional regulation, behavioral withdrawal, anxious/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behaviors, and aggression ^a	– Children with ASD were not included in the normative sample + Frequently used to examine co-occurring psychiatric conditions in ASD^b ± Various scales, such as the Withdrawn scale and the DSM-Pervasive Developmental Problems scale, have been shown to discriminate between ASD and typically developing peers and those with other psychiatric disorders; however, it shows decreased specificity for discriminating between ASD and other developmental disorders^c – Susceptible to responder bias as it relies on solely parent/caregiver reported information
Behavior Assessment System for Children, 3rd Edition,	BASC-3 PRS-P	2-5 years	Parent/caregiver, multidimensional behavior rating system, comprised of 138	+ Theoretical rationale behind clinical and adaptive scales corresponds to DSM-V diagnostic

Name	Abbreviation	Age Range	Description	Strengths/Weaknesses
Parent Rating Scales, Preschool Version			items aimed at assessing aspects of emotional and behavioral functioning in young children ^d ; Parents/caregivers rate items based on <i>frequency</i> of the behaviors.	criteria rather than statistically derived factor loadings, possibly enhancing its clinical utility^d + Updated standardization samples and norms (Released in 2015) + Inclusion of Developmental Social Disorders scale to capture symptomology of atypical development; Research on the utility of DSD scale has documented its ability to discriminate between typically developing children and those with ASD^e – Limited amount of research documenting and replicating BASC-3 PRS-P discriminative abilities to differentiate between ASD and other DDs^{e, f} – Susceptible to responder bias as it relies on solely parent/caregiver reported information

^aAchenbach and Rescorla (2000). ^bPandolfi, Magyar, and Dill (2009); Vaillancourt et al. (2017). ^cHavdahl et al. (2016); Muratori et al. (2011); Antonio Narzisi et al. (2013). ^dReynolds and Kamphaus (2015). ^eBradstreet et al. (2017); C. L. Myers et al. (2014). ^fJuechter (2012).

Table 2.

Summary of BASC-2 Research Findings

Citations	Number of Participants	Age Range	Gender (M: F)	BASC-2 Results
Kent (2006)	50 (AD=32, HFA=11, PDD-NOS=7)	8-18	43:7	- AD and HFA scored groups scored significantly higher on scales of: - DSD - Resiliency - No mean differences were found between genders across all content scales.
Volker et al. (2010)	124 (ASD=62, TD=62)	6-16	ASD =55:7	- DSD cut score of 60 accurately detected 95% of children with ASD from children who were TD - ASD group scored significantly higher on scales of: - Hyperactivity - Depression - Attention - Withdrawal - Atypicality - Anxiety
Mahan and Matson (2011)	80 (ASD=38; TD=42)	6-16	ASD =30:8	- ASD group scored significantly higher on scales of: - Hyperactivity - Conduct Problems - Depression

Citations	Number of Participants	Age Range	Gender (M: F)	BASC-2 Results
				○ Atypicality ○ Somatization ○ Attention Problems • ASD group scored significantly lower on all adaptive scales
Goldin et al. (2014)	151 (ASD=57, DD=28, TD=66)	2-16	ASD= 67:10; DD=19:9	• ASD group scored significantly higher than DD, TD groups on scales of: ○ Withdrawal ○ Atypicality • ASD group scored significantly lower than DD, TD groups on scales of: ○ Social skills ○ Leadership ○ Activities of Daily Living ○ Functional Communication
Juechter (2012)	158 (ASD=58, DD=28, TD=34)	25-37 months	ASD=43:15, DD= 14:14	• ASD group scored significantly higher than TD group on: ○ Atypicality ○ Withdrawal ○ Attention • No significant differences found between ASD and DD groups on any clinical or adaptive scales

Citations	Number of Participants	Age Range	Gender (M: F)	BASC-2 Results
				• Most useful DSD score for best distinguishing between ASD and DD groups= 60; this score produced sensitivity and specificity values of .78 and .69, respectively.
Myers, Gross, & McReynolds (2014)	156 (ASD=70; DD=86)	24-71 months	ASD= 59:11 DD=68:18	• ASD group scored significantly lower than DD group on scales of: ○ Social Skills ○ Functional Communication ○ Internalizing Problems ○ Externalizing Problems • DD group scored significantly higher than ASD group on scales of: ○ Hyperactivity ○ Aggression ○ Anxiety ○ Depression
Bradstreet et al. (2017)	224 (ASD=117; DD=55; TD=52)	24-63 months	ASD= 87:30 DD= 21:55	• DSD cut score of 61 differentiated ASD group from other groups (DD and TD), detecting 72 % of children with ASD and 63 % of children without ASD • ASD group scored significantly higher on scales of: ○ Atypicality ○ Withdrawal ○ Attention Problems • ASD group scored significantly lower on scales of:

Citations	Number of Participants	Age Range	Gender (M: F)	BASC-2 Results
				○ Adaptability ○ Social Skills ○ Daily Living Activities ○ Functional Communication ○ Emotional Self-Control ○ Resiliency
Gardner (2017)	232 (ASD only=79, ID=41, ASD/ID=122)	25-71 months	ASD=62:17, ID=35:6, ASD/ID=92:30	• DSD score of 60 produced sensitivity of .89 (all groups), but specificity of .15 (ASD vs. ID) • ASD only group scored significantly higher than ASD/ID and ID groups on the Withdrawal scale • ASD and ASD/ID groups scored significantly higher than ID only group and equally elevated on scales of: ○ Atypicality ○ Attention Problems • ASD and ASD/ID groups scored lower than ID only group on scales of: ○ Social Skills ○ Functional Communication • DSD scale scores were comparable for African American and White children in the referred sample

BIBLIOGRAPHY

- Achenbach, T. M., & Rescorla, L. A. (2000). Manual for the ASEBA preschool forms and profiles
In. Burlington, VT: University of Vermont, Research Center for Children, Youth,
& Families.
- Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment
of young children with Autism Spectrum Disorders. *Child Neuropsychol*, 12(4-5),
269-277. doi:10.1080/09297040500473714
- Aldinger, K. A., Lane, C. J., Veenstra-VanderWeele, J., & Levitt, P. (2015). Patterns of
Risk for Multiple Co-Occurring Medical Conditions Replicate Across Distinct
Cohorts of Children with Autism Spectrum Disorder. *Autism Res*, 8(6), 771-781.
doi:10.1002/aur.1492
- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental
disorders. In (5th ed.). Washington, DC: Author.
- Arndt, S., Turvey, C., & Andreasen, N. C. (1999). Correlating and predicting psychiatric
symptom ratings: Spearman's r versus Kendall's tau correlation. *J Psychiatr Res*,
33(2), 97-104. doi:10.1016/s0022-3956(98)90046-2
- Bagner, D. M., Rodriguez, G. M., Blake, C. A., Linares, D., & Carter, A. S. (2012).
Assessment of behavioral and emotional problems in infancy: a systematic
review. *Clin Child Fam Psychol Rev*, 15(2), 113-128. doi:10.1007/s10567-012-
0110-2
- Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., . . .
Dowling, N. F. (2018). Prevalence of Autism Spectrum Disorder Among Children

- Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveill Summ*, 67(6), 1-23.
doi:10.15585/mmwr.ss6706a1
- Barnard-Brak, L., Brewer, A., Chesnut, S., Richman, D., & Schaeffer, A. M. (2016). The sensitivity and specificity of the social communication questionnaire for autism spectrum with respect to age. *Autism Res*, 9(8), 838-845. doi:10.1002/aur.1584
- Bauman, M. L. (2010). Medical comorbidities in autism: challenges to diagnosis and treatment. *Neurotherapeutics*, 7(3), 320-327. doi:10.1016/j.nurt.2010.06.001
- Bishop, S. L., Farmer, C., & Thurm, A. (2015). Measurement of nonverbal IQ in autism spectrum disorder: scores in young adulthood compared to early childhood. *J Autism Dev Disord*, 45(4), 966-974. doi:10.1007/s10803-014-2250-3
- Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *Am J Intellect Dev Disabil*, 116(5), 331-343.
doi:10.1352/1944-7558-116.5.331
- Bishop, S. L., Luyster, R., Richler, J., & Lord, C. (2008). Diagnostic Assessment. In *Autism Spectrum Disorders in Infants and Toddlers: Diagnosis, Assessment, and Treatment* (pp. 23-49): Guilford Press.
- Bradstreet, L. E., Juechter, J. I., Kamphaus, R. W., Kerns, C. M., & Robins, D. L. (2017). Using the BASC-2 Parent Rating Scales to Screen for Autism Spectrum Disorder

- in Toddlers and Preschool-Aged Children. *J Abnorm Child Psychol*, 45(2), 359-370. doi:10.1007/s10802-016-0167-3
- Buie, T., Campbell, D. B., Fuchs, G. J., 3rd, Furuta, G. T., Levy, J., Vandewater, J., . . . Winter, H. (2010). Evaluation, diagnosis, and treatment of gastrointestinal disorders in individuals with ASDs: a consensus report. *Pediatrics*, 125 Suppl 1(Supplement 1), S1-18. doi:10.1542/peds.2009-1878C
- Burns, T. G., King, T. Z., & Spencer, K. S. (2013). Mullen Scales of Early Learning: The Utility in Assessing Children Diagnosed With Autism Spectrum Disorders, Cerebral Palsy, and Epilepsy. *Applied Neuropsychology: Child*, 2(1), 33-42. doi:10.1080/21622965.2012.682852
- Buros Center for Testing. (2014). Weisse review of Leiter-3. from Buros Institute, University of Nebraska-Lincoln. <https://marketplace.unl.edu/buros/>
- Carter, A. S., Black, D. O., Tewani, S., Connolly, C. E., Kadlec, M. B., & Tager-Flusberg, H. (2007). Sex Differences in Toddlers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 37(1), 86-97. doi:10.1007/s10803-006-0331-7
- Chawarska, K., Klin, A., & Volkmar, F. R. (2008). Autism Spectrum Disorders in Infants and Toddlers: Diagnosis, Assessment, and Treatment. *Guilford Publications*.
- Christensen, D. L., Bilder, D. A., Zahorodny, W., Pettygrove, S., Durkin, M. S., Fitzgerald, R. T., & Yeargin-Allsopp, M. (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and

- developmental disabilities monitoring network. *Journal of Developmental & Behavioral Pediatrics*, 37(1), 1-8.
- Christensen, D. L., Maenner, M. J., Bilder, D., Constantino, J. N., Daniels, J., Durkin, M. S., . . . Dietz, P. (2019). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 4 Years - Early Autism and Developmental Disabilities Monitoring Network, Seven Sites, United States, 2010, 2012, and 2014. *MMWR Surveill Summ*, 68(2), 1-19. doi:10.15585/mmwr.ss6802a1
- Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics*.
- Constantino, J. N., Zhang, Y. I., Frazier, T., Abbacchi, A. M., & Law, P. (2010). Sibling recurrence and the genetic epidemiology of autism. *American Journal of Psychiatry*, 167(11), 1349-1356.
- De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European child & adolescent psychiatry*, 7(3), 131-136.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26): Sage publications.
- Dickerson, A. S., Rahbar, M. H., Pearson, D. A., Kirby, R. S., Bakian, A. V., Bilder, D. A., . . . Slay Wingate, M. (2017). Autism spectrum disorder reporting in lower socioeconomic neighborhoods. *Autism*, 21(4), 470-480. doi:10.1177/1362361316650091

- Doshi-Velez, F., Ge, Y., & Kohane, I. (2014). Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. *Pediatrics*, *133*(1), e54-e63. doi:10.1542/peds.2013-0819
- Durkin, M. S., Maenner, M. J., Meaney, F. J., Levy, S. E., DiGuseppi, C., Nicholas, J. S., . . . Schieve, L. A. (2010). Socioeconomic inequality in the prevalence of autism spectrum disorder: evidence from a U.S. cross-sectional study. *PloS one*, *5*(7), e11551. doi:10.1371/journal.pone.0011551
- Elliott, C. (2007). Differential Ability Scales . San Antonio, TX: Harcourt Assessment.
- Canadian Journal of School Psychology*, *22*(1), 128-132.
- Farmer, C., Golden, C., & Thurm, A. (2016). Concurrent validity of the differential ability scales, second edition with the Mullen Scales of Early Learning in young children with and without neurodevelopmental disorders. *Child Neuropsychol*, *22*(5), 556-569. doi:10.1080/09297049.2015.1020775
- Feeley, K., & Jones, E. (2006). Addressing challenging behaviour in children with Down syndrome: The use of applied behaviour analysis for assessment and intervention. *Down Syndrome Research and Practice*, *11*(2), 64-77.
- Gardner, L. M., Campbell, J. M., Bush, A. J., & Murphy, L. (2017). Comparing Behavioral Profiles for Autism Spectrum Disorders and Intellectual Disabilities Using the BASC-2 Parent Rating Scales–Preschool Form. *Journal of Psychoeducational Assessment*, *36*(6), 535-551. doi:10.1177/0734282916689438

- Gardner, L. M., Campbell, J. M., Bush, A. J., & Murphy, L. (2018). Comparing Behavioral Profiles for Autism Spectrum Disorders and Intellectual Disabilities Using the BASC-2 Parent Rating Scales–Preschool Form. *Journal of Psychoeducational Assessment*, 36(6), 535-551. doi:10.1177/0734282916689438
- Giarelli, E., Wiggins, L. D., Rice, C. E., Levy, S. E., Kirby, R. S., Pinto-Martin, J., & Mandell, D. S. (2010). Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disabil Health J*, 3(2), 107-116. doi:10.1016/j.dhjo.2009.07.001
- Goldberg, W. A., Thorsen, K. L., Osann, K., & Spence, M. A. (2008). Use of home videotapes to confirm parental reports of regression in autism. *J Autism Dev Disord*, 38(6), 1136-1146. doi:10.1007/s10803-007-0498-6
- Goldin, R. L., Matson, J. L., Konst, M. J., & Adams, H. L. (2014). A comparison of children and adolescents with ASD, atypical development, and typical development on the Behavioral Assessment System for Children, (BASC-2). *Research in Autism Spectrum Disorders*, 8(8), 951-957.
- Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling Down the Wait Time: Exploring Models to Minimize the Delay from Initial Concern to Diagnosis and Treatment of Autism Spectrum Disorder. *Pediatric clinics of North America*, 63(5), 851-859. doi:10.1016/j.pcl.2016.06.007

- Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *J Autism Dev Disord*, 37(4), 613-627. doi:10.1007/s10803-006-0280-1
- Grondhuis, S. N., & Mulick, J. A. (2013). Comparison of the Leiter International Performance Scale—Revised and the Stanford-Binet Intelligence Scales, in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, 118(1), 44-54.
- Guli, L. A., Semrud-Clikeman, M., Lerner, M. D., & Britton, N. (2013). Social Competence Intervention Program (SCIP): A pilot study of a creative drama program for youth with social difficulties. *The Arts in Psychotherapy*, 40(1), 37-44. doi:<https://doi.org/10.1016/j.aip.2012.09.002>
- Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., . . . Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. *Mol Autism*, 6, 36. doi:10.1186/s13229-015-0019-y
- Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: an examination of developmental functioning, autistic symptoms, and coexisting behavior problems in toddlers. *Journal of Autism and Developmental Disorders*, 39(12), 1715.
- Hass, M. R., Brown, R. S., Brady, J., & Johnson, D. B. (2012). Validating the BASC-TRS for Use With Children and Adolescents With an Educational Diagnosis of

- Autism. *Remedial and Special Education*, 33(3), 173-183.
doi:10.1177/0741932510383160
- Havdahl, K. A., von Tetzchner, S., Huerta, M., Lord, C., & Bishop, S. L. (2016). Utility of the Child Behavior Checklist as a Screener for Autism Spectrum Disorder. *Autism Res*, 9(1), 33-42. doi:10.1002/aur.1515
- Hill, A. P., Zuckerman, K. E., Hagen, A. D., Kriz, D. J., Duvall, S. W., van Santen, J., . . . Fombonne, E. (2014). Aggressive Behavior Problems in Children with Autism Spectrum Disorders: Prevalence and Correlates in a Large Clinical Sample. *Res Autism Spectr Disord*, 8(9), 1121-1133. doi:10.1016/j.rasd.2014.05.006
- Holtmann, M., Bölte, S., & Poustka, F. (2007). Autism spectrum disorders: Sex differences in autistic behaviour domains and coexisting psychopathology. *Developmental Medicine & Child Neurology*, 49(5), 361-366. doi:10.1111/j.1469-8749.2007.00361.x
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398): John Wiley & Sons.
- Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. *Pediatric clinics of North America*, 59(1), 103-111, xi.
doi:10.1016/j.pcl.2011.10.018
- Hus, V., Gotham, K., & Lord, C. (2014). Standardizing ADOS domain scores: separating severity of social affect and restricted and repetitive behaviors. *J Autism Dev Disord*, 44(10), 2400-2412. doi:10.1007/s10803-012-1719-1

- Jansiewicz, E. M., Goldberg, M. C., Newschaffer, C. J., Denckla, M. B., Landa, R., & Mostofsky, S. H. (2006). Motor signs distinguish children with high functioning autism and Asperger's syndrome from controls. *Journal of Autism and Developmental Disorders*, 36(5), 613-621.
- Jo, H., Schieve, L. A., Rice, C. E., Yeargin-Allsopp, M., Tian, L. H., Blumberg, S. J., . . . Boyle, C. A. (2015). Age at autism spectrum disorder (ASD) diagnosis by race, ethnicity, and primary household language among children with special health care needs, United States, 2009–2010. *Maternal and Child Health Journal*, 19(8), 1687-1697.
- Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, 120(5), 1183-1215.
doi:10.1542/peds.2007-2361
- Jones, E. J., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: a review of prospective studies of infants at risk. *Neurosci Biobehav Rev*, 39, 1-33. doi:10.1016/j.neubiorev.2013.12.001
- Juechter, J. I. (2012). Differential diagnosis of autism spectrum disorders using the BASC-2 Parent Rating Scales Preschool Form.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous child*, 2(3), 217-250.

- Kent, E. (2006). Differential diagnosis of autism spectrum disorders using the Behavior Assessment System for Children--Parent rating scale (BASC-2 PRS) content scales.
- Khowaja, M., Robins, D. L., & Adamson, L. B. (2018). Utilizing two-tiered screening for early detection of autism spectrum disorder. *Autism*, 22(7), 881-890.
doi:10.1177/1362361317712649
- King, B. H. (2016). Psychiatric comorbidities in neurodevelopmental disorders. *Current Opinion Neurology*, 29(2), 113-117.
doi:10.1097/wco.000000000000029910.1097/WCO.0000000000000299.
- Knutsen, J., Crossman, M., Perrin, J., Shui, A., & Kuhlthau, K. (2019). Sex differences in restricted repetitive behaviors and interests in children with autism spectrum disorder: An Autism Treatment Network study. *Autism*, 23(4), 858-868.
doi:10.1177/1362361318786490
- Koegel, L. K., Koegel, R. L., Ashbaugh, K., & Bradshaw, J. (2014). The importance of early identification and intervention for children with or at risk for autism spectrum disorders. *International journal of speech-language pathology*, 16(1), 50-56.
- Kuravackel, G., & Ruble, L. (2014). Autism spectrum disorders: Reasons for treatment referrals across the developmental life span. *Journal of the Kentucky Medical Association*, 12(3), 51-56.

- Laerd Statistics. (2017). Binomial logistic regression using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from <https://statistics.laerd.com/>
- Landa, R. J., Gross, A. L., Stuart, E. A., & Faherty, A. (2013). Developmental trajectories in children with and without autism spectrum disorders: the first 3 years. *Child Dev*, 84(2), 429-442. doi:10.1111/j.1467-8624.2012.01870.x
- Lappe, M., Lau, L., Dudovitz, R. N., Nelson, B. B., Karp, E. A., & Kuo, A. A. (2018). The Diagnostic Odyssey of Autism Spectrum Disorder. *Pediatrics*, 141(Suppl 4), S272-S279. doi:10.1542/peds.2016-4300C10.1542/peds.2016-4300C.
- Lappé, M., Lau, L., Dudovitz, R. N., Nelson, B. B., Karp, E. A., & Kuo, A. A. (2018). The Diagnostic Odyssey of Autism Spectrum Disorder. *Pediatrics*, 141(Suppl 4), S272-S279. doi:10.1542/peds.2016-4300C
- Larsson, H. J., Eaton, W. W., Madsen, K. M., Vestergaard, M., Olesen, A. V., Agerbo, E., . . . Mortensen, P. B. (2005). Risk Factors for Autism: Perinatal Factors, Parental Psychiatric History, and Socioeconomic Status. *American Journal of Epidemiology*, 161(10), 916-925. doi:10.1093/aje/kwi123
- Lawson, L. P., Joshi, R., Barbaro, J., & Dissanayake, C. (2018). Gender Differences During Toddlerhood in Autism Spectrum Disorder: A Prospective Community-Based Longitudinal Follow-Up Study. *J Autism Dev Disord*, 48(8), 2619-2628. doi:10.1007/s10803-018-3516-y
- Le Couteur, A., Haden, G., Hammal, D., & McConachie, H. (2008). Diagnosing Autism Spectrum Disorders in Pre-school Children Using Two Standardised Assessment

- Instruments: The ADI-R and the ADOS. *Journal of Autism and Developmental Disorders*, 38(2), 362-372. doi:10.1007/s10803-007-0403-3
- Leaf, J. B., Taubman, M., Bloomfield, S., Palos-Rafuse, L., Leaf, R., McEachin, J., & Oppenheim, M. L. (2009). Increasing social skills and pro-social behavior for three children diagnosed with autism through the use of a teaching package. *Research in Autism Spectrum Disorders*, 3(1), 275-289.
doi:<https://doi.org/10.1016/j.rasd.2008.07.003>
- Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *J Autism Dev Disord*, 37(5), 894-910. doi:10.1007/s10803-006-0218-7
- Lever, A. G., & Geurts, H. M. (2016). Psychiatric Co-occurring Symptoms and Disorders in Young, Middle-Aged, and Older Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 46(6), 1916-1930.
doi:10.1007/s10803-016-2722-8
- Levy, S. E., Giarelli, E., Lee, L.-C., Schieve, L. A., Kirby, R. S., Cunniff, C., . . . Rice, C. E. (2010). Autism spectrum disorder and co-occurring developmental, psychiatric, and medical conditions among children in multiple populations of the United States. *Journal of Developmental & Behavioral Pediatrics*, 31(4), 267-275.
- Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C. M., & Fryer, G. E. (2008). Disparities in Diagnosis and Access to Health Services for Children with Autism: Data from the National Survey of Children's Health.

Journal of Developmental & Behavioral Pediatrics, 29(3), 152-160.

doi:10.1097/DBP.0b013e318165c7a0

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(6), 466-474.

doi:10.1016/j.jaac.2017.03.013

Lord, C., Corsello, C., & Grzadzinski, R. (2014). Diagnostic Instruments in Autistic Spectrum Disorders. In F. R. Volkmar, R. Paul, S. J. Rogers, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders* (Fourth ed.). Retrieved from

<https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118911389.hautc25>.

doi:10.1002/9781118911389.hautc25

Lord, C., Luyster, R., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) [Manual: Toddler Module]*. Torrance, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavorne, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) [Manual: Modules 1-4]*. Torrance, 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. *J Autism Dev Disord*, 24(5), 659-685.
- Lyall, K., Croen, L. A., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., . . . Newschaffer, C. (2017). The Changing Epidemiology of Autism Spectrum Disorders. *Annu Rev Public Health*, 38, 81-102. doi:10.1146/annurev-publhealth-031816-044318
- 10.1146/annurev-publhealth-031816-044318. Epub 2016 Dec 21.
- Mahan, S., & Matson, J. L. (2011). Children and adolescents with autism spectrum disorders compared to typically developing controls on the Behavioral Assessment System for Children, (BASC-2). *Research in Autism Spectrum Disorders*, 5(1), 119-125.
- Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuseppi, C., Durkin, M. S., . . . Kirby, R. S. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *Am J Public Health*, 99(3), 493-498. doi:10.2105/AJPH.2007.131243
- Mannion, A., & Leader, G. (2013). Comorbidity in autism spectrum disorder: A literature review. *Research in Autism Spectrum Disorders*, 7(12), 1595-1616.
- Marascuilo, L. A., & McSweeney, M. (1977). Nonparametric and distribution-free methods for the social sciences.

- Mari-Bauset, S., Zazpe, I., Mari-Sanchis, A., Llopis-Gonzalez, A., & Morales-Suarez-Varela, M. (2014). Food selectivity in autism spectrum disorders: a systematic review. *J Child Neurol*, 29(11), 1554-1561. doi:10.1177/0883073813498821
- Matheis, M., Matson, J. L., Hong, E., & Cervantes, P. E. (2018). Gender Differences and Similarities: Autism Symptomatology and Developmental Functioning in Young Children. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-018-3819-z
- McCrimmon, A. W., & Rostad, K. (2013). Test Review: Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part II): Toddler Module. *Journal of Psychoeducational Assessment*, 32(1), 88-92. doi:10.1177/0734282913490916
- Merrell, K. W. (2008). *Behavioral, social, and emotional assessment of children and adolescents* (3rd ed.). New York: Lawrence Erlbaum Associates.
- Ming, X., Brimacombe, M., Chaaban, J., Zimmerman-Bier, B., & Wagner, G. C. (2008). Autism spectrum disorders: concurrent clinical disorders. *Journal of child neurology*, 23(1), 6-13.
- Minshawi, N. F., Hurwitz, S., Morriss, D., & McDougale, C. J. (2015). Multidisciplinary assessment and treatment of self-injurious behavior in autism spectrum disorder and intellectual disability: integration of psychological and biological theory and approach. *J Autism Dev Disord*, 45(6), 1541-1568. doi:10.1007/s10803-014-2307-

- Moeschler, J. B. (2019). Neurodevelopmental Disabilities: Global Developmental Delay, Intellectual Disability, and Autism. In R. E. Pyeritz, B. R. Korf, & W. W. Grody (Eds.), *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics* (7th ed., pp. 61-79): Academic Press.
- Mosconi, M. W., Mohanty, S., Greene, R. K., Cook, E. H., Vaillancourt, D. E., & Sweeney, J. A. (2015). Feedforward and feedback motor control abnormalities implicate cerebellar dysfunctions in autism spectrum disorder. *Journal of Neuroscience*, 35(5).
- Mullen, E. M. (1995). Mullen Scales of Early Learning. In. Circle Pines, MN: AGS.
- Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry*, 27(5), 657-669.
- Muratori, F., Narzisi, A., Tancredi, R., Cosenza, A., Calugi, S., Saviozzi, I., . . . Calderoni, S. (2011). The CBCL 1.5–5 and the identification of preschoolers with autism in Italy. *Epidemiology and Psychiatric Sciences*, 20(4), 329-338.
- Myers, C. L., Bour, J. L., Sidebottom, K. J., Murphy, S. B., & Hakman, M. (2010). Same constructs, different results: Examining the consistency of two behavior-rating scales with referred preschoolers. *Psychology in the Schools*, 47(3), 205-216.
doi:10.1002/pits.20465

- Myers, C. L., Gross, A. D., & McReynolds, B. M. (2014). Broadband behavior rating scales as screeners for autism? *J Autism Dev Disord*, *44*(6), 1403-1413.
doi:10.1007/s10803-013-2004-7
- Myers, S. M., & Johnson, C. P. (2007). Management of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1162-1182. doi:10.1542/peds.2007-2362
- Narzisi, A., Calderoni, S., Maestro, S., Calugi, S., Mottes, E., & Muratori, F. (2013). Child Behavior Check List 1(1/2)-5 as a tool to identify toddlers with autism spectrum disorders: a case-control study. *Res Dev Disabil*, *34*(4), 1179-1189.
doi:10.1016/j.ridd.2012.12.020
- Narzisi, A., Calderoni, S., Maestro, S., Calugi, S., Mottes, E., & Muratori, F. (2013). Child Behavior Check List 1½-5 as a tool to identify toddlers with autism spectrum disorders: a case-control study. *Research in Developmental Disabilities*, *34*(4), 1179-1189.
- Neumeyer, A. M., Anixt, J., Chan, J., Perrin, J. M., Murray, D., Coury, D. L., . . . Parker, R. A. (2018). Identifying Associations Among Co-Occurring Medical Conditions in Children With Autism Spectrum Disorders. *Acad Pediatr*.
doi:10.1016/j.acap.2018.06.014
- Nguyen, C. T., Krakowiak, P., Hansen, R., Hertz-Picciotto, I., & Angkustsiri, K. (2016). Sociodemographic Disparities in Intervention Service Utilization in Families of Children with Autism Spectrum Disorder. *J Autism Dev Disord*, *46*(12), 3729-3738. doi:10.1007/s10803-016-2913-3

- Ozonoff, S., Gangi, D., Hanzel, E. P., Hill, A., Hill, M. M., Miller, M., . . . Iosif, A. M. (2018). Onset patterns in autism: Variation across informants, methods, and timing. *Autism Res, 11*(5), 788-797. doi:10.1002/aur.1943
- Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *J Clin Child Adolesc Psychol, 34*(3), 523-540. doi:10.1207/s15374424jccp3403_8
- Ozonoff, S., & Iosif, A. M. (2019). Changing conceptualizations of regression: What prospective studies reveal about the onset of autism spectrum disorder. *Neurosci Biobehav Rev, 100*, 296-304. doi:10.1016/j.neubiorev.2019.03.012
- Ozonoff, S., Iosif, A. M., Young, G. S., Hepburn, S., Thompson, M., Colombi, C., . . . Rogers, S. J. (2011). Onset patterns in autism: correspondence between home video and parent report. *J Am Acad Child Adolesc Psychiatry, 50*(8), 796-806 e791. doi:10.1016/j.jaac.2011.03.012
- Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M., & Rogers, S. J. (2008). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism, 12*(5), 457-472. doi:10.1177/1362361308096402
- Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., . . . Iosif, A. M. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: a baby siblings research consortium study. *J Child Psychol Psychiatry, 56*(9), 988-998. doi:10.1111/jcpp.12421

- Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the child behavior checklist 1.5-5 in a sample of children with autism spectrum disorders. *J Autism Dev Disord*, 39(7), 986-995. doi:10.1007/s10803-009-0716-5
- Peters-Scheffer, N., Didden, R., Mulders, M., & Korzilius, H. (2010). Low intensity behavioral treatment supplementing preschool services for young children with autism spectrum disorders and severe to mild intellectual disability. *Res Dev Disabil*, 31(6), 1678-1684. doi:10.1016/j.ridd.2010.04.008
- Powell, K. K., Heymann, P., Tsatsanis, K. D., & Chawarska, K. (2018). Assessment and Diagnosis of Infants and Toddlers with Autism Spectrum Disorder. In S. Goldstein & S. Ozonoff (Eds.), *Assessment of autism spectrum disorders* (Second Edition ed., pp. 96-129). New York, NY: The Guilford Press.
- Reinhardt, V. P., Wetherby, A. M., Schatschneider, C., & Lord, C. (2015). Examination of sex differences in a large sample of young children with autism spectrum disorder and typical development. *J Autism Dev Disord*, 45(3), 697-706. doi:10.1007/s10803-014-2223-6
- Rescorla, L. A., Winder-Patel, B. M., Paterson, S. J., Pandey, J., Wolff, J. J., Schultz, R. T., & Piven, J. (2017). Autism spectrum disorder screening with the CBCL/1(1/2)-5: Findings for young children at high risk for autism spectrum disorder. *Autism*, 1362361317718482. doi:10.1177/1362361317718482
- Reynolds, C. R., & Kamphaus, R. W. (2015). BASC-3 behavioral and emotional screening system manual. *Circle Pines, MN: Pearson*.

- Richards, C., Jones, C., Groves, L., Moss, J., & Oliver, C. (2015). Prevalence of autism spectrum disorder phenomenology in genetic disorders: a systematic review and meta-analysis. *Lancet Psychiatry*, 2(10), 909-916. doi:10.1016/S2215-0366(15)00376-4
- Richler, J., Bishop, S. L., Kleinke, J. R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *J Autism Dev Disord*, 37(1), 73-85. doi:10.1007/s10803-006-0332-6
- Robins, D. L., Casagrande, K., Barton, M., Chen, C. M., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, 133(1), 37-45. doi:10.1542/peds.2013-1813
- Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). Leiter international performance scale,(Leiter-3). *Los Angeles: Western Psychological Services*.
- Ros-Demarize, R., Bradley, C., Kanne, S. M., Warren, Z., Boan, A., Lajonchere, C., . . . Carpenter, L. A. (2020). ASD symptoms in toddlers and preschoolers: An examination of sex differences. *Autism Res*, 13(1), 157-166. doi:10.1002/aur.2241
- Rutter, M., Bailey, A., & Lord, C. (2003). The Social Communication Questionnaire. In. Los Angeles, CA: Western Psychological Services.
- Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism Diagnostic Interview-Revised. In. Los Angeles, CA: Western Psychological Services.

- Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, 10(1), 91-103.
doi:10.1007/bf02408436
- Schopler, E., Van Bourgondien, M., Wellman, J., & Love, S. (2010). Childhood autism rating scale—Second edition (CARS2): Manual. *Los Angeles: Western Psychological Services*.
- Shapiro, E. S., & Heick, P. F. (2004). School psychologist assessment practices in the evaluation of students referred for social/behavioral/emotional problems. *Psychology in the Schools*, 41(5), 551-561. doi:10.1002/pits.10176
- Shumway, S., Thurm, A., Swedo, S. E., Deprey, L., Barnett, L. A., Amaral, D. G., . . . Ozonoff, S. (2011). Brief report: symptom onset patterns and functional outcomes in young children with autism spectrum disorders. *J Autism Dev Disord*, 41(12), 1727-1732. doi:10.1007/s10803-011-1203-3
- Sigman, M., Dijamco, A., Gratier, M., & Rozga, A. (2004). Early detection of core deficits in autism. *Mental retardation and developmental disabilities research reviews*, 10(4), 221-233.
- Sipes, M., Matson, J. L., Worley, J. A., & Kozlowski, A. M. (2011). Gender differences in symptoms of Autism Spectrum Disorders in toddlers. *Research in Autism Spectrum Disorders*, 5(4), 1465-1470. doi:10.1016/j.rasd.2011.02.007

- Soke, G. N., Maenner, M. J., Christensen, D., Kurzius-Spencer, M., & Schieve, L. A. (2018a). Prevalence of Co-occurring Medical and Behavioral Conditions/Symptoms Among 4- and 8-Year-Old Children with Autism Spectrum Disorder in Selected Areas of the United States in 2010. *J Autism Dev Disord*, 48(8), 2663-2676. doi:10.1007/s10803-018-3521-1
- Soke, G. N., Philofsky, A., Diguseppi, C., Lezotte, D., Rogers, S., & Hepburn, S. (2011). Longitudinal changes in Scores on the Autism Diagnostic Interview--Revised (ADI-R) in pre-school children with autism: Implications for diagnostic classification and symptom stability. *Autism*, 15(5), 545-562. doi:10.1177/1362361309358332
- Spreckley, M., & Boyd, R. (2009). Efficacy of applied behavioral intervention in preschool children with autism for improving cognitive, language, and adaptive behavior: a systematic review and meta-analysis. *J Pediatr*, 154(3), 338-344. doi:10.1016/j.jpeds.2008.09.012
- Stephens, R. L., Langworthy, B., Short, S. J., Goldman, B. D., Girault, J. B., Fine, J. P., . . . Gilmore, J. H. (2018). Verbal and nonverbal predictors of executive function in early childhood. *J Cogn Dev*, 19(2), 182-200. doi:10.1080/15248372.2018.1439493
- Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997). Motor imitation in young children with autism: what's the object? *J Abnorm Child Psychol*, 25(6), 475-485. doi:10.1023/a:1022685731726

- Szatmari, P., Chawarska, K., Dawson, G., Georgiades, S., Landa, R., Lord, C., . . . Halladay, A. (2016). Prospective Longitudinal Studies of Infant Siblings of Children With Autism: Lessons Learned and Future Directions. *J Am Acad Child Adolesc Psychiatry*, 55(3), 179-187. doi:10.1016/j.jaac.2015.12.014
- Szatmari, P., Georgiades, S., Duku, E., Bennett, T. A., Bryson, S., Fombonne, E., . . . Pathways in, A. S. D. S. T. (2015). Developmental trajectories of symptom severity and adaptive functioning in an inception cohort of preschool children with autism spectrum disorder. *JAMA Psychiatry*, 72(3), 276-283. doi:10.1001/jamapsychiatry.2014.2463
- Thomson, K., Burnham Riosa, P., & Weiss, J. A. (2015). Brief Report of Preliminary Outcomes of an Emotion Regulation Intervention for Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 45(11), 3487-3495. doi:10.1007/s10803-015-2446-1
- Vaillancourt, T., Haltigan, J. D., Smith, I., Zwaigenbaum, L., Szatmari, P., Fombonne, E., . . . Bryson, S. (2017). Joint trajectories of internalizing and externalizing problems in preschool children with autism spectrum disorder. *Dev Psychopathol*, 29(1), 203-214. doi:10.1017/S0954579416000043
- Volker, M. A., Lopata, C., Smerbeck, A. M., Knoll, V. A., Thomeer, M. L., Toomey, J. A., & Rodgers, J. D. (2010). BASC-2 PRS Profiles for Students with High-Functioning Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 40(2), 188-199. doi:10.1007/s10803-009-0849-6

Volkmar, F. R., Siegel, M., Woodbury-Smith, M., King, B., McCracken, J., State, M., . . .

Adolescent Psychiatry Committee on Quality, I. (2014). Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*, 53(2), 237-257.

doi:10.1016/j.jaac.2013.10.013

Walrath, R. (2011). Behavior Rating Scale. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of Child Behavior and Development* (pp. 228-228). Boston, MA: Springer US.

Werner, E., & Dawson, G. (2005). Validation of the phenomenon of autistic regression using home videotapes. *Arch Gen Psychiatry*, 62(8), 889-895.

doi:10.1001/archpsyc.62.8.889

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief report: Recognition of autism spectrum disorder before one year of age: a retrospective study based on home videotapes. *J Autism Dev Disord*, 30(2), 157-162.

Wetherby, A. M., Watt, N., Morgan, L., & Shumway, S. (2007). Social Communication Profiles of Children with Autism Spectrum Disorders Late in the Second Year of Life. *Journal of Autism and Developmental Disorders*, 37(5), 960-975.

doi:10.1007/s10803-006-0237-4

White, S. P., Weitlauf, A. S., & Warren, Z. E. (2012). Early Diagnosis of Autism Spectrum Disorder: Progress, Challenges, and Remaining Questions for Families

- and Professionals. In R. M. Hodapp (Ed.), *International Review of Research in Developmental Disabilities Volume 43* (Vol. 43, pp. 47-85): Academic Press.
- White, S. W., Keonig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: A review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858-1868.
- Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115(5), 654-657. doi:10.1161/CIRCULATIONAHA.105.594929
- Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Fein, D., Kasari, C., Pierce, K., . . . Wetherby, A. (2015a). Early Identification and Interventions for Autism Spectrum Disorder: Executive Summary. *Pediatrics*, 136 Suppl 1(Supplement 1), S1-9. doi:10.1542/peds.2014-3667B
- Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., . . . Natowicz, M. R. (2015b). Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research. *Pediatrics*, 136 Suppl 1(Supplement 1), S60-81. doi:10.1542/peds.2014-3667E
- Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., . . . Wagner, S. (2015c). Early Screening of Autism Spectrum Disorder: Recommendations for Practice and Research. *Pediatrics*, 136 Suppl 1(Supplement 1), S41-59. doi:10.1542/peds.2014-3667D

- Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of autism spectrum disorders. *Behav Brain Res*, 251, 133-146.
doi:10.1016/j.bbr.2013.04.004
- Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., . . . Yirmiya, N. (2009). Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics*, 123(5), 1383-1391. doi:10.1542/peds.2008-1606
- Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *Int J Dev Neurosci*, 23(2-3), 143-152. doi:10.1016/j.ijdevneu.2004.05.001
- Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Roberts, W., Szatmari, P., . . . Vaillancourt, T. (2016). Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort. *Autism Res*, 9(7), 790-800. doi:10.1002/aur.1585