Differential Diagnosis of Autism Spectrum Disorders Using the BASC-2 Parent Rating Scales Preschool Form

Julia I. Juechter
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/cps_diss

Recommended Citation
doi: https://doi.org/10.57709/3089580

This Dissertation is brought to you for free and open access by the Department of Counseling and Psychological Services at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Counseling and Psychological Services Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
DIFFERENTIAL DIAGNOSIS OF AUTISM SPECTRUM DISORDERS USING THE BASC-2 PARENT RATING SCALES PRESCHOOL FORM

Julia I. Juechter
ABSTRACT

DIFFERENTIAL DIAGNOSIS OF AUTISM SPECTRUM DISORDERS AND OTHER DEVELOPMENTAL DELAYS USING THE BASC-2 PARENT RATING SCALES-PRESCHOOL FORM

by

Julia I. Juechter

The Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus 2004) is a behavior rating scale commonly used in preschool settings. In addition to measuring behavioral constructs such as hyperactivity, social skills, and adaptive functioning, the BASC-2 includes a Developmental Social Disorders (DSD) content scale that evaluates the presence of behaviors commonly associated with pervasive developmental disorders, including items related to self-stimulation, withdrawal and poor socialization. This study compared the T-scores of toddler and preschool-aged children diagnosed with an autism spectrum disorder (ASD) to children diagnosed with other developmental delays, and typically developing children using the BASC-2 Parent Rating Scales, Preschool form. Participants from the ASD group obtained significantly higher T-scores than the typically developing group on the Hyperactivity, Atypicality, Withdrawal, and Attention Problems scales, and obtained significantly lower T-scores on the Adaptability, Social Skills, Activities of Daily Living, and and Functional Communication scales. Significant differences were not observed between participants in the ASD group and those diagnosed with other developmental delays. However, the DSD scale was effective in distinguishing between groups, with participants in the ASD group obtaining significantly higher T-scores on the DSD scale than those diagnosed with other developmental delays and typically developing children.
DIFFERENTIAL DIAGNOSIS OF AUTISM SPECTRUM DISORDERS AND OTHER DEVELOPMENTAL DELAYS USING THE BASC-2 PARENT RATING SCALES-PRESCHOOL FORM
by
Julia I. Juechter

A Dissertation

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

Atlanta, GA 2012
ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Randy Kamphaus and Dr. Diana Robins, for the mentorship and support throughout this process. Thank you for believing in me and allowing me to pursue research that I have a passionate for. You have transformed me from a consumer of research to an actual researcher, and helped me to grow both personally and professionally.

I would also like to thank my other committee member, Dr. Bridget Dever and Dr. Gregory Brack, for their behind the scenes support since this endeavor began, and for stepping up to the table when I needed you. This work would not have been completed without your contributions.

Finally, I would like to thank my mother, sister, and friends for all of the encouragement that you all have offered me not only in these final months of research, but throughout my academic career. What matters to me most is not who calls me “doctor” at the end of this long process, but who still calls me “friend.”
# TABLE OF CONTENTS

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

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION AND REVIEW OF THE LITERATURE</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Purpose of the Study</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Diangostic Classifications of Developmental Disorders</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Differential Diagnosis of Developmental Disabilities</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Research Questions</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>METHODS</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Procedures Utilized for Research</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Instruments</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Data Analyses</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>RESULTS</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>DISCUSSION</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Implications</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Future Research</td>
<td>79</td>
</tr>
</tbody>
</table>

References ........................................................................................................ 80

Appendix ............................................................................................................. 100
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive Statistics for Demographic Variables for Total Sample and Diagnostic Categories</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>Correlation Matrix of Demographic Variables by BASC-2 Clinical Scale Scores for Total Sample</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>Correlation Matrix of Demographic Variables by BASC-2 Adaptive Scale Scores and DSD Scale Score for Total Sample</td>
<td>44</td>
</tr>
<tr>
<td>4</td>
<td>MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (Autism and PD-NOS groups)</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (Autism and PD-NOS groups)</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (GDD and Lang Delay groups)</td>
<td>47</td>
</tr>
<tr>
<td>7</td>
<td>MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (GDD and Lang Delay groups)</td>
<td>48</td>
</tr>
<tr>
<td>8</td>
<td>MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (M-CHAT groups)</td>
<td>49</td>
</tr>
<tr>
<td>9</td>
<td>MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (M-CHAT groups)</td>
<td>51</td>
</tr>
<tr>
<td>10</td>
<td>MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (M-CHAT Typical and Norm Group)</td>
<td>52</td>
</tr>
<tr>
<td>11</td>
<td>MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (M-CHAT Typical and Norm Group)</td>
<td>53</td>
</tr>
<tr>
<td>12</td>
<td>MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (ASDs Collapsed)</td>
<td>54</td>
</tr>
<tr>
<td>13</td>
<td>MANCOVA for BASC-2 Adaptive Scales Using Age (months) as Covariate (ASDs Collapsed)</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (ASDs Separated)</td>
<td>57</td>
</tr>
<tr>
<td>15</td>
<td>MANCOVA for BASC-2 Adaptive Scales Using Age (months) as Covariate (ASDs Separated)</td>
<td>59</td>
</tr>
<tr>
<td>16</td>
<td>ANOVA for BASC-2 Developmental Social Disorders Scale (ASDs Collapsed)</td>
<td>61</td>
</tr>
<tr>
<td>17</td>
<td>ANOVA for BASC-2 Developmental Social Disorders Scale (ASDs Separated)</td>
<td>62</td>
</tr>
<tr>
<td>18</td>
<td>Sensitivity and Specificity of the DSD Scale in Identifying ASD Versus Non-ASD</td>
<td>63</td>
</tr>
<tr>
<td>19</td>
<td>Correlation Matrix of BASC-2 DSD Scale T-Score with ADOS and CARS Total Scores</td>
<td>64</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean ( T )-scores of clinical, adaptive, and DSD scales by diagnostic classification group</td>
<td>69</td>
</tr>
<tr>
<td>2</td>
<td>Mean ( T )-scores of clinical, adaptive, and DSD scales by study population</td>
<td>69</td>
</tr>
<tr>
<td>3</td>
<td>Scatterplot of DSD ( T )-Scores by diagnostic classification group</td>
<td>73</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>Autism Spectrum Disorder</td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>Autistic Disorder</td>
<td></td>
</tr>
<tr>
<td>PDD-NOS</td>
<td>Pervasive Developmental Disorder-Not Otherwise Specified</td>
<td></td>
</tr>
<tr>
<td>Other DD</td>
<td>Other Developmental Delay</td>
<td></td>
</tr>
<tr>
<td>TD</td>
<td>Typically Developing</td>
<td></td>
</tr>
<tr>
<td>GDD</td>
<td>Global Developmental Delay</td>
<td></td>
</tr>
<tr>
<td>DLD</td>
<td>Developmental Language Delay</td>
<td></td>
</tr>
<tr>
<td>DSD</td>
<td>Developmental Social Disorders</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION AND REVIEW OF THE LITERATURE

Developmental disabilities affect approximately 5 to 10% of children under the age of five. Conservative estimates of the prevalence of all Autism Spectrum Disorders (ASDs) is 1/110, with the prevalence rate of Autistic Disorder (AD) accounting for approximately 22/10,000 (Center for Disease Control and Prevention, 2009; Saracino et al., 2010). However, a recent surveillance study conducted by the Autism and Developmental Disabilities Monitoring (ADDM) Network found that the overall prevalence rate for ASDs had increased to an average of 1 in 88 children across sites involved in the study (CDC, 2012). It is difficult to ascertain a reliable prevalence rate for autism and related Pervasive Developmental Disorders (PDD), as the definitions used for the different disorders vary between clinicians and research teams (Fombonne, 2003). Adding further confusion, diagnostic criteria for PDDs are vague and do not specify how much impairment (or competence) is needed for diagnosis (Towbin, 2005).

Although some studies have noted that ASDs can be difficult to distinguish from global developmental delay in young children (e.g., Lord, 1995), research now supports the diagnosis of ASDs in children as young as two years of age (Boyd, Odom, Humphreys, & Sam, 2010; Kleinman et al., 2008; Ventola et al., 2007; Woods & Wetherby, 2003). Further, diagnosis of ASDs in this age group have been found to be relatively stable over time (Kleinman et al., 2008), and research indicates that intervention provided before 42 months of age has greater impact than intervention that is
started after the age of five (Crane & Winsler, 2008; Wetherby et al., 2004; Woods & Weatherby, 2003).

Mandell and colleagues (2007; 2005) point out that despite increasing evidence that autism can be accurately diagnosed in toddler and preschool-age children, many children often go undiagnosed until they reach school-age. They also note that many children are initially misdiagnosed, as symptoms of autism are mistaken for Attention Deficit Hyperactivity Disorder (ADHD), language impairment, obsessive-compulsive disorder (OCD), or oppositional defiant disorder (ODD). Although these disorders may co-occur with autism, it is important that autism be diagnosed as well so that appropriate interventions can be implemented. Additionally, the presence of comorbid behavior disorders may require specific interventions. Challenging behaviors such as aggression interfere with learning, contribute to the use of physical restraint and use of medication, and predict placement in more restrictive learning environments (Mandell, Ittenbach, Levy, & Pinto-Martin, 2007; Mandell, Novak, & Zubritsky, 2005). Therefore, it is important to assess and treat challenging behaviors in children with ASDs.

Behavior rating scales are commonly used in the assessment of autism spectrum disorders as measures of co-occurring behavior problems as well as deficits in social and adaptive behaviors. School-based clinicians frequently employ behavior rating scales when conducting screenings and evaluations for a variety of referral concerns, including developmental delays (Volker et al., 2010). Rating scales can also be a valuable tool for clinicians, such as school psychologists, to help guide decision making when evaluating children with a possible ASD. As these clinicians, and many others, may not have training in autism-specific assessment, behavior rating scales that clinicians are familiar
with, such as the BASC-2 (Reynolds & Kamphaus, 2004), may help to highlight behavioral deficits indicative of an ASD, and aid clinicians in making appropriate referrals for further evaluation and intervention. The purpose of the current study is to investigate the utility of the BASC-2 Parent Rating Scales Preschool Form as a diagnostic tool in the assessment of toddler and preschool age children at risk for autism spectrum disorders.

**Diagnostic Classification of Developmental Disorders**

Autism spectrum disorders (ASDs) are neurodevelopmental disorders that are characterized by impairments in communication and social interactions, and may include the presence of stereotyped or repetitive behaviors or circumscribed interests (American Psychological Association, 2004). Conceptualization of the autism spectrum suggests that it exists on a continuum of impairment, with Autistic Disorder (AD) being most severe presentation, and Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) representing the less severe end of the spectrum (Volkmar et al., 1994). Despite the lack of reliable diagnostic criteria for PDD-NOS, it is one of the most commonly diagnosed PDDs. However, the presentation of autistic symptoms in children diagnosed with PDD-NOS varies greatly, making reliable use of the diagnostic category difficult (Chlebowski et al., 2010).

ASDs also can be difficult to differentiate from other developmental disorders, particularly in younger populations. Because a deficit in communication is one of the defining features of autistic disorder, children with language delays or global delays of functioning may present as at-risk for ASDs (Ventola et al., 2007). Additionally, some symptoms from the DSM-IV-TR diagnostic criteria may be developmentally
inappropriate for toddlers. The following is a review of the current diagnostic criteria for autistic disorder, Pervasive Developmental Disorder-Not Otherwise Specified, global developmental delay, and developmental language delays.

**Autism**

**Primary features and diagnostic criteria.** The American Psychiatric Association (2004) and the ICD-10 (WHO, 1992) describe Autistic Disorder as a pervasive developmental disorder characterized by abnormal or impaired development in communication, social interaction, and a restricted range of interests or repetitive behaviors. DSM-IV-TR criteria for Autistic Disorder (presented in Table 1) require the presence of at least six symptoms, with at least two symptoms from the social domain, one or more symptoms from the communication domain, and at least one symptom from the restricted, repetitive, stereotyped behaviors domain, and stipulates that symptoms must be evident prior to the third birthday (APA, 2004).

**Developmental course of autism.** Approximately 30% to 54% of parents of children diagnosed with autism report having had concerns before their child's first birthday (De Giacomo & Fombonne, 1998; Volkmann, Stier, & Cohen, 1985), and 80% to 90% of parents indicate abnormal development before 24 months (De Giacomo & Fombonne, 1998). These are conservative estimates given that they are based primarily upon retrospective accounts by parents, and thus limited by parents' lay knowledge of child development and subjectivity (Chawarska & Volkmar, 2005). Other studies (e.g., Maestro et al., 2001) utilize videotapes (i.e., home movies) of children who were later diagnosed with autism to analyze behavioral deficits and abnormalities in the first few years of life.
Limited eye contact, lack of motor imitation and babbling, and decreased social responsiveness have all been noted in infants who are later diagnosed with autism (Dawson, Meltzoff, & Kuhl, 2000; Klin et al., 2004; Sparling, 2001). Infants also may exhibit excessive startle responses, arousal regulation difficulties, sleep difficulties, unusual sensitivity to touch and other sensori stimuli, and motor problems (Dawson et al., 2000; Sparling, 2001). When compared with a developmentally delayed control group matched for both mental age and chronological age, children with autism were identified as being less likely to anticipate being picked up, show affection towards familiar people, show interest in children other than siblings, reach for a familiar person, or play simple interaction games such as peek-a-boo in the year before their first birthday (Klin, Volkmar, & Sparrow, 1992). Parents of young children later diagnosed with autism described them as being either extremely difficult (e.g., agitated, difficult to soothe, unable to get on a consistent feeding and sleep schedule) or extremely passive as infants (Rogers & DiLalla, 1990). In studies comparing videotapes of infants 6 months old and younger who were later diagnosed with autism with those of typically developing babies, infants with autism were less likely to attend to people visually, smile at others, or vocalize (Maestro et al., 2002). However, in a similar study of infants aged 8 to 10 months, smiling, vocalizations, and looking at others occurred at similar rates when comparing children with autism and typically developing controls. The only distinguishing behavior between the two groups at this age was that infants with autism were less likely to respond to their name being called in comparison with typically developing infants (Werner et al., 2000). More recent prospective studies of high-risk infants (those with an older sibling with an ASD) have also shown that gaze behavior and
reported parental concerns at 6 months of age did not predict a later diagnosis of autism (Ozonoff et al., 2008; Young, Merin, Rogers, & Ozonoff, 2009). However, between 12 and 18 months of age, parental report of developmental concerns such as atypicalities in visual tracking, motor mannerisms, social-communication behavior, and language do reflect differences between children later diagnosed with autism and typically developing children (Ozonoff et al., 2008; Ozonoff et al., 2010; Zwaigenbaum et al., 2009). Other studies have noted that in addition to being less responsive to the sound of their own name, children with autism are less responsive to visual stimuli at this age than children with intellectual disabilities or typical controls (Baranek, 1999; Osterling, Dawson, & Munson, 2002).

Many parents of children with autism begin to recognize abnormalities in their child's development between their second and third birthdays. Delays in language acquisition, loss of skills such as speech or eye contact, and emerging stereotyped mannerisms most often prompt parents to seek medical or psychological advice regarding their children's development (De Giacomo & Fombonne, 1998; Rogers & DiLalla, 1990; Short & Schopler, 1988). Between 20% and 40% of children with autism are estimated to experience regression in acquired skills between 12 and 24 months of age, such as a loss of language or motor skills (Fombonne & Chakrabarti, 2001; Lord, Shulman, & DiLavore, 2004; Rogers & DeLalla, 1990; Tuchman & Rapin, 1997). Common symptoms observed at this age include limited imitation, not looking at others, limited facial expressions, preference for being alone, failure to point, sometimes replaced with using another person's hand as a tool, failure to use instrumental gestures (such as raising their arms to be picked up), and atypical sensory responses (i.e., hyper or hyposensitivity
to sounds, textures, taste, and visual stimuli). Stereotyped movements (e.g., finger flicking mannerisms), unusual vocalizations, and unusual sensory interests (e.g., preoccupation with fans or lights) begin to emerge between the second and third years of life. Children with autism also demonstrate lack of interest in other children, limited functional play, and no evidence of pretend play at this age (Chawarska and Volkmar, 2005; Lord, 1995; Lord & Pickles, 1996; Stone, Hoffman, Lewis, & Ousley, 1994; Wimpory, Hobson, Williams, & Nash, 2000).

**Pervasive Developmental Disorder-Not Otherwise Specified**

A diagnosis of Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) is frequently given to those individuals who exhibit severe deficits in social interaction in conjunction with impaired communication skills and/or stereotyped behaviors, interests, and activities, but who do not meet the criteria for another PDD due to atypical, subthreshold, or late onset symptomology (APA, 2004; Buitelaar & van der Gaag, 1998; Towbin, 2005). These criteria most closely correspond to that of atypical autism as defined in the ICD-10, but may also fall under the ICD-10 classification of Other Pervasive Developmental Disorder or Pervasive Developmental Disorder, Unspecified, depending upon the symptomology (WHO, 1992). Many clinicians conceptualize PDD-NOS as a “milder” form of autism. For example, children with PDD-NOS may be able to perceive emotional characteristics, but not apply them (Towbin, 2005). Children may receive a diagnosis of PDD-NOS with or without exhibiting deficits in receptive or expressive language; however, individuals without social deficits but who exhibit language impairments and stereotyped movements should not be diagnosed with PDD-NOS.
Differential diagnosis of PDD-NOS is often complex, particularly when distinguishing the disorder from Autistic Disorder or other developmental delays, as the level of impairment necessary to obtain a diagnosis of PDD-NOS is still undefined. Other conditions to be considered in the differential diagnosis include intellectual disabilities, language disorders, avoidant disorder, anxiety disorders, reactive attachment disorder, and schizo-affective disorders (Towbin, 2005). In some cases these diagnoses may co-occur with PDD-NOS, whereas in others a diagnosis of PDD-NOS may be sufficient to explain symptoms exhibited by the individual. Additionally, many children diagnosed with PDD-NOS exhibit symptoms of inattention similar to children diagnosed with ADHD. In a study by Luteijn et. al. (2000) comparing children diagnosed with PDD-NOS to those with a diagnosis with ADHD, results revealed no significant differences between the PDD-NOS group and the ADHD group on general psychopathology, general autistic symptomology, social and self-help skills, and attention problems. However, children with PDD-NOS were significantly more withdrawn than children in the ADHD group. The authors further suggested that co-morbid diagnoses of PDD-NOS and ADHD be examined and permitted in future revisions to the DSM.

The category of PDD-NOS as described in the DSM-IV-TR is often criticized for its lack of well-defined criteria, leading to diverse and possibly inaccurate interpretations and diagnoses (Buitelaar & van der Gaag, 1998; Luteijn et. al., 2000 Towbin, 2005). More distinctive diagnostic criteria for PDD-NOS are essential for communication between clinicians, therapists, and families to ensure proper treatment planning. Further, the lack of more explicit criteria for PDD-NOS impedes research related to prevalence, etiology, and prognosis (Buitelaar & van der Gaag, 1998; Towbin, 2005).
Global Developmental Delay

Global Developmental Delay (GDD) is not a disorder defined in the DSM-IV-TR or the ICD-10; however, this diagnosis is commonly used by medical practitioners and psychologists alike. This term is typically reserved for use in younger children (ages 2 to 5), whereas Mental Retardation, a formal DSM-IV diagnosis (now referred to in the professional community as Intellectual Disability), is more commonly applied when working with older children when results of IQ tests are more valid and reliable. GDD is operationally defined as significant delay (i.e., two or more standard deviations below the mean on norm-referenced developmental testing) in at least two or more domains: motor skills, cognition, speech and language, personal/social skills (including play and recreation), or self-help skills (Shevell, 1998; Shevell et. al., 2003; Shevell, Majnemer, Platt, Webster, & Birnbaum, 2005). These children experience limitations in their overall rate of learning and acquisition of skills when compared with other children of the same chronological age. Children with GDD also have a high frequency of hearing and vision impairment (20% to 50%). In addition, seizure disorders, behavioral problems, sleep disturbances, and feeding problems frequently co-occur with GDD and need to be addressed (Shevell, 1998). Outcomes for children initially diagnosed with GDD include deficits in academics (i.e., math, reading, and handwriting skills), social skills, and behavior; and children with GDD fare significantly worse than children diagnosed with language impairment (Shevell et. al., 2005). Early diagnosis of developmental delay is imperative to improve the outcomes for these children (Shevell et. al., 2003).

Estimates of the prevalence of GDD in children younger than 5 years range from 1% to 3% (compared to a prevalence rate of 1% for mental retardation), although the
underlying causes of GDD often go undetermined (Boyle, Yeargin-Allsop, Doernberg, Holmgreen, Murphy, & Schendel, 1996). Previously identified etiologies include cerebral dysgenesis, hypoxic-ischemic encephalopathy (HIE), chromosomal abnormalities, psychosocial neglect, and in utero or antenatal exposure to toxins (Shevell et al. 2001; Tervo, 2006). As Shevell and colleagues (2001) point out, more than half of children diagnosed with GDD have a theoretically preventable etiology (HIE, exposure to toxins, and psychosocial neglect). Environmental factors such as culture, parenting skills, and opportunity may also affect the phenotype and diagnosis of GDD (Shevell et. al., 2003). Although primary care physicians are becoming more likely to refer cases of developmental delay to specialists such as child neurologists and developmental pediatricians for the recommended standardized assessments (Shevell, 1998; Shevell, Majnemer, Rosenbaum, & Abrahamowicz, 2001), GDD is typically diagnosed based on clinical impressions of a child's overall functioning during a single visit to the pediatrician (Shevell et. al., 2005).

**Developmental Language Delay**

Speech and language disorders are the most common developmental disorders seen in preschool children (Tervo, 2007). Developmental Language Delay refers to a delay in developmentally appropriate use of communicative expressive and/or receptive language skills, in the absence of cognitive impairment or hearing loss (Shevell et. al., 2005). Prevelance estimates of developmental language disorders are approximately 7.4% for kindergarten-age children; however, children with language delays may fall below the clinical cut-off for impairment at variable periods at different times throughout the course of their lives (Toppelberg & Shapiro, 2000). Speech and language delays often
first present in the toddler years, with most parents reporting initial concerns between the ages of 18 and 23 months (Tervo, 2007). Although expressive language skills are slower to develop than receptive language skills, toddlers and preschoolers often present with mixed receptive-expressive language delays. Children with language delays often exhibit impaired social skills, leading to poor social-emotional functioning and behavioral problems, particularly in preschool aged children (Rescorla, Ross, & McClure, 2007; Tervo, 2007). Because deficits in communication skills as well as social interaction are also common features of PDDs, differential diagnosis can be difficult.

Differential Diagnosis of Developmental Disabilities

Accurate diagnosis of autism spectrum disorders (ASD) requires a multimethod approach that includes observation of the child, caregiver interview, assessment of developmental levels, detailed developmental history, and screening of associated disorders (Filipek et al., 2000). Currently, two instruments, the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999) and the Autism Diagnostic Interview-Revised (ADI-R: Rutter et al., 2003) are recognized as the “gold standard” of autism assessment. However, both of these measures are time-intensive and require extensive training prior to administration. Following is a review of commonly used instruments in the assessment of autism spectrum disorders in young children.

Measures of ASD Symptomology

Autism Diagnostic Observation Schedule. The Autism Diagnostic Observation Schedule (ADOS; Lord et al. 1999) is a semi-structured assessment of social interaction, communication, and play designed to measure symptoms of autism in children and adults. The ADOS contains four modules, which are designed for use according to the
developmental and language level of the individual. Modules 1 and 2 are most commonly used with pre-school aged children: Module 1 is appropriate for children with only single words or no speech, and Module 2 is intended for use with young children with phrase speech. Although coded, repetitive behaviors and stereotyped interests are not included in the scoring algorithm. The authors report good inter-rater reliability of items \((k \geq 0.6)\), and that disagreements between raters most often occurred when differentiating between diagnoses of autistic disorder and PDD-NOS. When discriminating between autism and non-autism, Lord et al. (2000) reported sensitivities of 1.00 and .95 for modules 1 and 2, respectively; and sensitivities of .94 and .89 when differentiating PDD-NOS from non-spectrum cases.

**Autism Diagnostic Interview-Revised.** The Autism Diagnostic Interview-Revised (ADI-R: Rutter et al., 2003) is a semi-structured interview used to assess autistic symptomology across three domains: communication, reciprocal social interaction, and repetitive behavior and stereotyped patterns. During the interview, the informant provides information based on current behavior, as well as behavior observed when the child was 4 to 5 years old. The authors of the ADI-R report good inter-rater reliability, with kappas ranging from .62 to .89, and good internal consistency. However, Lord et al. (1993) reported that the ADI-R did not discriminate well between children with autism with a mental age below 18 months and children without autism who have cognitive or language impairments. Additionally, researchers report that the ADI-R identifies significantly fewer toddlers with ASD when the behavioral domain is included in the algorithm (Saemundsen et al. 2003; Ventola et al. 2006; Wiggins & Robins, 2008).
Although the ADI-R has been validated in studies with children over the age of 4, follow-up studies indicate that ADI-R diagnoses given before the age of three are less stable (Charman & Baird, 2002; Chawarski et al. 2007). Charman et al. (2005) monitored the ADI-R classifications of 26 children at four different time points between 2 and 7 years of age, and found that six participants changed diagnostic classification twice over the course of the study, while an additional eight participants changed diagnostic classifications at least once. Moss, Magiati, Charman, and Howlin (2008) examined the stability of ADI-R diagnoses for 35 children aged 22 to 54 months at the time of initial diagnosis. The mean age of the sample was 3.5 years at time 1, and 10.5 years at follow-up. All 35 participants scored above the cut-off for autism on all three domains of the ADI-R at time 1. At follow-up, 20% of children in the sample no longer met the algorithm cut-off for autism.

**Childhood Autism Rating Scale.** The *Childhood Autism Rating Scale* (CARS; Schopler, Reichler, & Renner, 1988) is a behavior rating scale designed to help differentiate autism from other developmental disorders. Trained observers rate the severity of children’s behaviors using 15 items, which include ratings of verbal and nonverbal communication, social and emotional responses, and restrictive and repetitive use of objects. Scores from individual items are added together to create a “Total Score” which assess the presence and severity of autism (i.e., no autism, mild to moderate autism, or severe autism). Authors of the CARS report good internal consistency (alpha = .94), high inter-rater reliability (.55 to .93), and good test-retest reliability (.88, r=.94) for the Total Score (Schopler et al. 1988).
Diagnostic validity of the CARS has been supported in the literature. Eaves and Milner (1993) found that the CARS accurately classified 47 of 48 subjects diagnosed with autism. Similarly, in a sample of children aged 18 months to 11 years, 54 out of 54 children diagnosed with autistic disorder in their sample received scores on the CARS in the mild-moderate autism or severe autism range (Rellini, Tortolani, Trillo, Carbone, & Montecchi, 2004), However, the four individuals diagnosed with PDD-NOS received scores below the clinical cut-off for autism. Consistent with these results, Perry et al. (2005) reported sensitivity of .88 for the CARS when DSM-IV criteria of autistic disorder (AD) was employed, whereas Ventola et al. (2006) found that the CARS had a sensitivity of .89 when used to diagnose ASD, and .96 for diagnoses of AD utilizing DSM-IV-TR criteria.

Perry et al. (2005) also found that children diagnosed with AD obtained significantly higher scores on the CARS than those diagnosed with PDD-NOS; however, the mean total score obtained by the PDD-NOS group was below the CARS clinical cut-off for autism. Chlebowski, Green, Barton, and Fein (2010) obtained similar results in a sample of two and four-year-olds, reporting that the AD group obtained significantly higher scores than the PDD-NOS group. In the two-year-old sample, lowering the clinical cut-off from 30 to 25.5 yielded a sensitivity of .93 and specificity of .91 for distinguishing ASD (i.e., autistic disorder or PDD-NOS) from non-ASD (e.g., language delay or global delay) or no diagnosis. In the four-year-old sample, a cut-off score of 25.5 resulted in 86% agreement between the CARS and DSM-IV ASD diagnoses (Cheblowski et al., 2010).
**Gilliam Autism Rating Scale, Second Edition.** The *Gilliam Autism Rating Scale, Second Edition* (GARS-2; Gilliam, 2006) is a norm-referenced measure designed to assess symptoms of autism in individuals aged three to 22 years. Like the original GARS, the GARS-2 contains three conceptually derived subscales, Stereotyped Behaviors, Communication, and Social Interaction, which are purported to assess the three core areas of impairment in autism as defined by the DSM-IV-TR (APA, 2000). Sums of scores from the three subscales are combined to determine the Autism Index (AI). In addition to the AI, a Parent Interview is included in the instrument to assess for the presence of developmental delays or abnormalities in socialization prior to age 3; however, this information does not contribute to the AI.

Studies of the psychometric properties of the original GARS (Gilliam, 1995) produced mixed results (Pandolfi, Magyar, & Dill, 2010). Mazefsky and Oswald (2006) found that the overall score (then referred to as the Autism Quotient, or AQ) underestimated the likelihood of autism, demonstrating lower than expected correlations with subscales of the ADOS and the ADI-R. Other researchers have obtained conflicting results regarding the sensitivity of the GARS, with sensitivity of the AQ ranging from low to adequate, with children diagnosed with ASDs other than Autistic Disorder (e.g., PDD-NOS) frequently obtaining scores below the cut-off for Autism on the GARS (South et al., 2002; Eaves et al., 2006; Lecavalier, 2005).

The factor structure of the GARS has also been called into question, with findings suggesting that the instrument placed more emphasis on the presence of stereotyped and repetitive behaviors than deficits in communication and social skills (Lecavalier, 2005). Additionally, exploratory and confirmatory factor analyses of the GARS-2 conducted by
Pandolfi et al. (2010) were not supportive of the three-factor structure, indicating that each of the conceptually derived subscales measured more than one construct, which the researchers suggest limit their interpretation. Additionally, many items on the GARS-2 are “double-barreled,” requiring raters to evaluate more than one behavior within the context of a single item (e.g., “Eats specific foods and refuses to eat what most people will usually eat”), which complicate psychometric analyses (Pandolfi et al., 2010, p. 1127). Finally, because of the continued emphasis on stereotyped and repetitive behaviors, Pandolfi and colleagues (2010) suggest that the GARS-2 may not adequately assess individuals with high functioning or milder presentations of autism.

**Social Communication Questionnaire.** The Social Communication Questionnaire (SCQ; Rutter et al. 2003) is a questionnaire based on revised algorithms of the ADI (Le Couteur et al., 1989). The SCQ contains 40 items that assess behaviors in the areas of reciprocal social interaction, language and communication, and repetitive and stereotyped patterns of behavior. The SCQ was developed with a sample of 200 children and adults (aged 4 to 40 years) previously assessed with the ADI or ADI-R, and is intended for use with individuals over the age of four. Receiver Operating Characteristic Curve Analyses produced and area under the curve of .86 when differentiating between ASD versus non-ASD, .92 when discriminating autism from intellectual disabilities, and .74 when differentiating between autistic disorder and other ASDs (Rutter et al. 2003).

Several studies have examined the diagnostic validity of the SCQ. In a sample of 4 to 18-year-olds for whom diagnosis was already known, researchers found the sensitivity and specificity to be 85% and 67%, respectively, for differentiating ASDs from other developmental delays (Berument, Rutter, Lork, Pickles, & Bailey, 1999).
Witwer and Lecavalier (2007) found that using a cut-off score of 15 yielded a sensitivity of .92 and specificity of .62 in a sample of children aged 4 to 14. Chandler et al. (2007) also found that the SCQ demonstrated adequate sensitivity (.88) and specificity (.72) when distinguishing between ASD and non-ASD school-aged children. However, in a sample of children aged 4 to 6, Eaves, Wingert, and Ho (2006) found that estimates of sensitivity and specificity were lower (.74 and .54, respectively).

Allen, Silove, Williams, and Hutchins (2006) investigated the sensitivity and specificity of the SCQ in a population of 2 to 6-year-olds at-risk for pervasive developmental disorders. They found that using a cut score of 11, the SCQ demonstrated good sensitivity (1.00) but low specificity (.58) for children ages 3 to 5. However, the SCQ performed poorly with children in the 2 to 3-year-old age group, rendering high rates of false positives. Snow and Lecavalier (2008) found even lower agreement between diagnosis and risk classification in a sample of toddlers and preschoolers. In discriminating between children with and without PDDs, the sensitivity of the SCQ was .70, and specificity was .52. However, Wiggins, Bakeman, Adamson, and Robins (2007) found the SCQ demonstrated adequate sensitivity (.89) and specificity (.89) in a sample of 37 children aged 17 to 45 months when a cut off score of 11 was used. Other studies using the SCQ with young children yielded estimates of sensitivity ranging from .61 to .76, and specificity ranging from .41 to .81 (Norris & Lecavalier, 2010; Osterling et al., 2010; Snow & Lecavalier, 2008). Researchers (e.g., Allen, Silove, Williams, & Hutchins, 2007; Norris & Lecavalier, 2010; Osterling et al. 2010) propose that whereas the SCQ may be useful as a screening instrument for ASD in children over the age of four, it is not appropriate for use with younger children (i.e., children aged 2 to 3 years).
**Autism Behavior Checklist.** The *Autism Behavior Checklist* (ABC; Krug, Arick, & Almond, 1980; 1993) is a screening instrument designed to measure autistic behaviors in individuals. The 57 items are divided into five subscales: sensory, relating, language, social and self-help, and body and object use. Items are weighted differently in accordance with degree to which the behavior is characterized as a symptom of autism. The authors of the ABC reported a split-half reliability of .94 for the Total Score; however, Eaves, Campbell, & Chambers (2000) assert that the inclusion of 100 typical individuals resulted in inflated reliability estimates. Construct validity was assessed by comparing the mean ABC scores of children with autism to four nonautistic groups: severely mentally retarded, severely emotionally disturbed, and typically developing children (Krug et al. 1980). Although all F ratios were statistically significant (p < .001), Eaves et al. (2000) point out that these results should be interpreted with caution because the same participants were used in the standardization sample of the ABC.

Examining estimates of internal consistency, Sturmey, Matson, and Sevin (1992) reported a coefficient alpha of .87 for the total score. In a sample of 107 children with pervasive developmental disorders, Eaves, Campbell, and Chambers (2000) found that using the lower cut-off score recommended by the authors resulted in overall classification accuracy of 80%, with a sensitivity of 77% and specificity of 91%. Evidence of inter-rater reliability varies, with correlations ranging from .08 to .72, the lowest agreement occurring when parent and teacher ratings are compared (Eaves et al. 2000, Volkmar et al. 1988). Studies of concurrent validity have also been contradictory, with validity coefficients between the ABC and the CARS ranging from .27 to .67 (Eaves & Milner, 1993; Sevin, Matson, Coe, Fee, & Sevin, 1991). Rellini, Tortolani, Trillo,
Carbone, and Montecchi (2004) found that the ABC did not distinguish children with autism from other developmental disorders such as language delay as well as the CARS. Whereas the CARS correctly classified 100% of individuals with Autistic Disorder, the ABC only classified 54% of cases correctly.

**Autism Spectrum Quotient: Children’s Version.** The *Autism Spectrum Quotient: Children’s Version* (AQ-Child; AuYeung, Baron-Cohen, Wheelwright, & Allison, 2007) is a parent-rated questionnaire designed to assess severity of autistic symptoms in children ages 4 to 11 years old. The AQ-Child was adapted from the previously published adolescent (AQ-Adol; Baron-Cohen, Hoekstra, Knickmeyer, & Wheelwright, 2006) and adult (AQ-Adult; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) versions of the scale, which were designed to measure of presence of traits and behaviors associated with autism in adults and adolescents with normal intelligence. Original items from the adult and adolescent questionnaires were revised to be developmental appropriate and related to children. Like the AQ-Adult and AQ-Adol versions, the AQ-Child assesses five areas domains of traits and behaviors associated with autism: social skills, attention to detail, attention switching, communication, and imagination. Each domain contains 10 statements that are rated on a 4-point Likert scale. Item scores are summed to produce the Autism Quotient (AQ), with higher scores representing the presence of more behaviors commonly associated with autism.

Original analyses of the AQ-Adult demonstrated that adults with an autism spectrum diagnosis scored significantly higher than those from the general population. 80% of individuals with a diagnosis of Asperger’s Syndrome or High Functioning Autism scored at or above 32 (the cut-off established by the authors), while only 2% of
adults in the control group scored at or above this cut-off. The authors reported high inter-rater and test-retest reliability (Baron-Cohen et al. 2001). For the AQ-Adol, adolescents diagnosed with an autism spectrum disorder also scored significantly higher than a group of matched controls (Baron-Cohen et al. 2006). Results from the adolescent sample were even more promising, with 90% of adolescents diagnosed with an ASD scoring at or above the cut-off of 30. 0% of adolescents in the control group obtained scores above the clinical cut-off.

Auyeung et al. (2007) conducted an exploratory study of the AQ-Child comparing children aged 7 to 11 years diagnosed with autistic disorder (AD; n=192) or Asperger’s Syndrome/high-functioning autism (AS/HFA; n=348) according to DSM-IV criteria, to a control group (n=1225). Children with a diagnosis of PDD-NOS were not included in the clinical sample. Results demonstrated satisfactory internal consistency for each of the five subscales (social skills = 0.93; attention to detail = 0.83; attention switching = 0.89; communication = 0.92; and imagination = 0.88). Test-retest reliability was also good (r = 0.85). Group differences on the were examined using an ANOVA; post-hoc tests revealed that both the AD and HFA clinical groups obtained significantly higher scores than the control group on each of the 5 subscales, but did not differ significantly from each other. ROC analyses showed a total item score of 76 (out of 150) demonstrated both high sensitivity (95%) and high specificity (95%).

Although the AQ-Child appears to demonstrate adequate internal consistency and test-retest reliability, as well as good sensitivity and specificity, the authors caution that several items concern behaviors that require the use of language. Therefore, they recommended that the AQ-Child be used with children with some speech, and with
intelligence in the borderline to low average range (i.e., $\geq 70$) or above (Auyeung et al. 2007).

**Pervasive Developmental Disorder Rating Scale.** The *Pervasive Developmental Disorder Rating Scale* (PDDRS; Eaves, 1993) is an adaptation of a rating scale developed to assess constructs across three subscales: arousal (aloneness, sensory stimulation, fascination for objects), affect (aggression, anxiety, fear, affect), and cognition (language, skill development) (Eaves, Campbell, & Chambers, 2000). The scale contains 51 items designed to assess behavioral characteristics of autism as described by the literature and the DSM-III-TR (American Psychological Association, 1987). Raters (i.e., teachers, parents, or other individuals who know the child well) judge the degree to which the child exhibits the behavior described using a 5-point Likert scale (Eaves, 1993).

The PDDRS was normed on 500 children diagnosed with pervasive developmental disorders, with 436 children in the sample reported to have a diagnosis of autistic disorder. Estimates of internal consistency were good, yielding the following reliability coefficients: Arousal $r = .90$, Affect $r = .84$, Cognition $r = .79$, and PDDRS Total $r = .92$. Test-retest reliability coefficients using the same raters were also adequate (Arousal $r = .89$, Affect $r = .87$, Cognition $r = .87$, PDDRS Total $r = .91$); however, test-retest reliability coefficients using two different raters (e.g., parent and teacher) were much lower (Arousal $r = .53$, Affect $r = .40$, Cognition $r = .44$, and PDDRS Total $r = .48$), thus demonstrating poor inter-rater reliability as well.

In a sample of 107 participants aged 4 to 11 years, Eaves et al. (2000) found that using a standard score at or above 85 for both the total score and the arousal score resulted in a sensitivity of 88% and a specificity of 88%. When compared with the ABC,
the PDDRS was found to measure similar constructs, and diagnostic agreement between
the two measures was 85%. Limitations of findings exist, however, as the authors did not
have knowledge of the criteria used to diagnose PDDs in the clinical sample.

**Multidimensional Behavior Rating Scales**

While the ADOS and the ADI-R are often considered the “gold standard” of
autism assessment, the high cost and time intensity of the specialized training required to
administer these instruments often limits their use in primary care and educational
settings. Researchers (e.g., Paul et al., 2004; Sikora, Hall, Hartley, Gerrard-Morris, &
Cagle, 2008; Stone, Ousley, Hepburn, Hogan, & Brown, 1999; Volker et al., 2010) point
out that parent-completed behavior rating scales are commonly used in both of these
settings, and may offer an alternative approach to screening for ASDs. Two commonly
used assessments of behavior in preschoolers include the Child Behavior Checklist
(CBCL; Achenbach & Rescorla, 2000) and the Behavior Assessment System for

**Child Behavior Checklist (CBCL).** The CBCL (Achenbach & Rescorla, 2000) is
a caregiver-completed measure of child behavior problems. There are two forms of the
CBCL: one for use with children ages 18 months to 5 years, 11 months, and the other for
children ages 6 to 18 years. Caregivers are asked to rate the degree to which each item is
true about their child’s behavior on a scale of 0 (not true) to 2 (very true or often true).
Scores are then summed and converted to T-scores to form seven syndrome scales
(Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Sleep
Problems, Attention Problems, and Aggressive Behavior), and two broadband scales:
“Internalizing” problems and “Externalizing” problems. The authors of the CBCL report
good test-retest reliability ($r = .68$ to 92, mean $r = .84$) and cross-informant agreement 
(mean mother–father $r = .61$, mean parent–child care provider $r = .65$). Validity results, 
from a sample from mental health and special education facilities and matched subjects 
from a normative sample, indicated that the internalizing and externalizing scales 
correctly classified 74% of referred children (Achenbach & Rescorla, 2000).

Luteijn et al. (2000) compared the CBCL profiles of children ages 5-12 diagnosed 
with PDD-NOS, ADHD, and co-morbid PDD and ADHD. Results indicated that the co-
morbid PDD/ADHD had significant higher Total Problem scores on the CBCL than the 
ADHD only group; however, the PDD-NOS only group did not differ significantly from 
the PDD/ADHD group or the ADHD only group on the Total Problems scale. Children in 
both the PDD-NOS and the PDD/ADHD group received significantly higher scores ($p < 
.001$) on the Social Problems scale than the ADHD only group. Children in the PDD-
NOS group also had significantly higher scores ($p < .001$) on the Withdrawn subscale 
than both the PDD/ADHD group and the ADHD only group. Interestingly, the PDD/NOS 
group did not differ significantly from the ADHD only group or the co-morbid 
PDD/ADHD group on the Attention Problems scale; however, the PDD/ADHD group 
obtained significantly higher scores on Attention Problems than the ADHD only group, 
leading the authors to speculate that social and attention problems interact in a negative 
way (Luteijn et al., 2000).

Sikora et al. (2008) examined the CBCL and GARS scores for 147 children ages 
36 to 71 months in comparison with autism classifications (Autism, ASD, Non-
Spectrum) based on scores obtained on the ADOS-G. Utilizing a cut-off score of $T \geq 70$, 
findings from the study indicated that the Withdrawn scale and the Pervasive
Developmental Problems scale on the CBCL accurately differentiated children with an ADOS classification of Autism from children with an ADOS classification of Non-Spectrum. The established cut-off for the Autism Quotient (≥ 90) obtained on the GARS did not significantly differentiate these groups.

Pandolfi, Magyar, and Dill (2009) conducted a confirmatory factor analysis of the CBCL 1.5-5 with 123 preschoolers diagnosed with ASD, and found that the CBCL measures the same constructs in children with an ASD as it does in the general population. Preschoolers in the ASD group also obtained significantly higher raw scores than the normative sample on each of the domain and syndrome scales with the exception of the Anxious/Depressed scale.

**Behavior Assessment System for Children, Second Edition--Parent Rating Scales (BASC-2 PRS).** The BASC-2 (Reynolds & Kamphaus, 2004) is a multidimensional assessment system that evaluates both clinical and adaptive aspects of behavior and emotional functioning. Parents rate the presence of behaviors in children on a four-point frequency scale (i.e., 0=Never, 1=Sometimes, 2=Often, and 3=Almost Always). Item raw scores are summed and converted into standardized T scores with a mean of 50 and a standard deviation of 10 for interpretation. For the clinical scales, higher scores represent more problematic behaviors, with T scores between 60 and 69 considered “at-risk”, and T scores of 70 or above being clinically significant. On the adaptive scales, lower scores are indicative of deficits, with T scores between 31 and 40 falling in the at-risk range and scores equal to or less than 30 considered clinically significant.
The BASC-2 PRS-Preschool Version (PRS-R; ages 2-5) reports T-scores for behaviors on the following scales: Hyperactivity, Aggression, Anxiety, Depression, Atypicality, Withdrawal, Somatization, Attention Problems, Adaptability, Social Skills, Activities of Daily Living, and Functional Communication. Additional content scales reported in the BASC-2 include Anger Control, Bullying, Developmental Social Disorders, Emotional Self-Control, Executive Functioning, Negative Emotionality, and Resiliency. Formulated via theoretical and empirical approaches, the content scales are derived from items belonging to both the primary scales listed previously, and items not on the primary clinical scales. The content scales combine items from multiple constructs to detect patterns of behavior. For example, the Developmental Social Disorders scale measures behaviors related to social skills, communication, and interests and activities (Reynolds & Kamphaus, 2004). Items contributing to the Developmental Social Disorders scale are presented in Table 2.

Authors of the BASC-2 PRS-P report adequate reliability and validity. Individual scales of the BASC-2 PRS-P have a median test-retest reliability of 0.77, and median inter-rater reliability of .74. The Behavioral Symptoms Index of the BASC-2 PRS-P is also highly correlated with the Achenbach System of Empirically Based Assessment Child Behavior Checklist for Ages 1.5-5 Total Problems score (0.75).

Mean T-scores for a clinical populations of children Pervasive Developmental Disorders are reported in the manual for the BASC-2; however, subtypes of PDD are not distinguished (i.e., Autism, Asperger's, PDD-NOS), nor are individual results of the preschool sample reported. The authors of the BASC-2 report mean T-scores in the clinically significant range for children and adolescents diagnosed with a PDD in the areas of
Atypicality, Withdrawal, and Functional Communication, and in the at-risk range for Hyperactivity, Attention Problems, Adaptability, Social Skills, Leadership, and Activities of Daily Living (Reynolds & Kamphaus, 2004). Children and adolescents from the clinical sample of those with Intellectual Disability and/or Developmental Delay obtained mean T-scores in the at-risk range for Atypicality, and only the child sample reported scores barely in the at-risk range for Withdrawal. While still in the at-risk range, scores on each of the the Adaptive scales for the Developmentally Delayed sample were significantly higher than those of the PDD clinical groups.

Kent (2006) conducted a study comparing content scale scores on the BASC-2-PRS of 50 children ages 8 to 18 who had a diagnosis of High Functioning Autism (HFA), Asperger’s Disorder (AD), and Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS). The author found that the Anger Control, Developmental Social Disorder, Emotional Self-Control, Executive Functioning, Negative Emotionality, and Resiliency content scales were highly correlated with DSM criteria for the various Pervasive Developmental Disorders. Kent (2006) found that the AD and HFA subgroups obtained mean T-scores in the clinically significant range (i.e., T \geq 70) on the Developmental Social Disorder and Resiliency content scales, while the mean T-scores of the PDD-NOS subgroup did not reach clinical significance on any of the content scales.

Volker et al. (2010) compared BASC-2-PRS scores between children with high functioning autism spectrum disorders (HFASD) and typically developing (TD) children aged 6 to 16 years (M=9.74, SD=2.22). Participants included 62 children with HFASDs and 62 TD children who were matched on age, gender, and ethnicity. Utilizing ANOVA and ANCOVA (adding demographics as a covariant), Volker et al. (2010) found
statistically significant differences ($p<.001$) between the HFASD group and the TD group across all four BASC composite scales (i.e., Externalizing Problems, Internalizing Problems, Adaptive Skills, and the Behavioral Symptoms Index). They also found statistically significant differences between the two groups among all individual PRS clinical and adaptive scales with the exception of Conduct Problems and Somatization. Volker et al. also compared cut-scores of 60, 65, and 70 on the Developmental Social Disorders content scale of the BASC-2 PRS, and found that using a cut score of 60 accurately screened 98% of participants known to have an ASD and screened out 95% of typically developing children.

Mahan and Matson (2011) recently conducted a study comparing the BASC-2-PRS profiles of 80 children and adolescents with ASD (n=38) and typically developing (TD) children (n=42). In contrast to the study done by Volker et al. (2010), the ASD group in this study was a heterogenous sample of children with both high and low-functioning autism spectrum disorders. Utilizing Mann-Whitney exact tests for each subscale, composite, and index score of the BASC-2-PRS, the researchers found that children in the ASD group obtained significantly higher scores on the externalizing composite scale and the behavioral symptoms index than the TD group. Children in the ASD group also scored significantly higher than the TD group on the hyperactivity, conduct problems, depression, somatization, atypicality, withdrawal, and attention problems subscales. The ASD group scored significantly lower on the overall adaptive composite and adaptive subscales including adaptability, social skills, leadership, activities of daily living, and functional communication (Mahan & Matson, 2011).
Research Questions and Hypotheses

The purpose of the current study was to examine the utility of the BASC-2 Parent Rating Scales- Preschool Form (BASC-2 PRS-P) for differentiating between children with different developmental delays. Specifically, this research addressed the following questions:

1. Do young children with autism spectrum disorders obtain significantly different scores on the clinical scales of the BASC-2 PRS-P compared to children with other developmental delays and typically developing children?

2. Does the BASC-2 DSD scale demonstrate adequate sensitivity and specificity for classifying toddlers and preschoolers with ASDs in comparison to children with other developmental delays and typically developing children?

3. Does the BASC-2 DSD scale demonstrate concurrent validity with empirically validated measures of ASD?

The first research question investigates the differences in BASC-2 PRS-P scores between children diagnosed with ASDs and other developmental delays (Other DDs). Specifically, this question examines the differences between toddlers and preschoolers diagnosed with AD, PDD-NOS, Other DDs (global delay and language delay), and typically developing (TD) children. Based on research conducted by Volker et al. (2010) and Mahan and Matson (2011), it is hypothesized that children with ASDs will exhibit significantly more impairment than TD children on all clinical and adaptive scales of the BASC-2 PRS-P. However, research using pilot data conducted by the current author (Juechter, Robins, Kamphaus, & Fein, 2011) indicated that whereas children with AD obtained significantly higher scores on the DSD scale of the BASC-2 than those with
PDD-NOS and other DDs, most scales of the BASC-2, including the DSD scale, did not distinguish between children diagnosed with PDD-NOS and those diagnosed with other DDs. This trend is expected to continue in the current study.

The second research question seeks to determine the sensitivity and specificity of the BASC-2 PRS-P DSD scale for classifying toddlers and preschoolers with ASDs. It is hypothesized that the DSD scale will demonstrate adequate sensitivity and specificity when distinguishing between children with ASDs (when the AD and PDD-NOS group are combined) and those not diagnosed with an ASD (i.e., other DDs and typically developing children).

The last research question seeks to establish concurrent validity between the DSD scale of the BASC-2 PRS-P and empirically validated measures of autism. It is hypothesized that the DSD content scale score will exhibit high agreement with scores obtained by participants on the ADOS and the CARS.
CHAPTER 2: METHODS

Sample Recruitment

Study participants were recruited from an ongoing screening study for developmental delays at the University of Connecticut and Georgia State University. Children between the ages of 14 and 30 months were screened at their pediatrician's office or early intervention provider using the Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al. 1999a; Robins et al. 2001) or the Modified Checklist for Autism in Toddlers-Revised (M-CHAT-R; Robins, Fein, & Barton, 2009). Children who initially failed the M-CHAT or M-CHAT-R (hereafter referred to as the M-CHAT, given that the version of the screening tool is not primary to the research questions in the current study) and subsequently received a clinical diagnosis of autism or other developmental delay (i.e., PDD-NOS, Developmental Language Delay, or Global Developmental Delay) after a comprehensive evaluation were identified for inclusion in the clinical groups. Children in the typically developing (TD) group were derived from a random sample of children who screened negative on the M-CHAT follow-up interview, and also screened negative on a follow-up screener, the Screening Tool for Autism in Two-Year-Olds (STAT: Stone & Ousley, 1997). Children who screened negative on the MCHAT follow-up interview but scored at-risk on the STAT were invited to participate in the full evaluation, and considered for inclusion in the clinical sample pending diagnosis. Children who participated in the full evaluation and were classified with as typically developing were also included in the TD sample. Children who were given other
diagnoses (e.g., ADHD, Developmental Coordination Disorder) were not included in the clinical or typical samples.

**Sample Characteristics**

BASC-2 scores for 158 children were examined in the study. Of the participants from the M-CHAT study for whom completed BASC-2 forms were received, 23 were diagnosed with Autism, 35 were diagnosed with PDD-NOS, 28 were diagnosed with other developmental delays (i.e., 20 participants diagnosed with Global Developmental Delay and 8 participants diagnosed with Language Delay), and 34 were determined to be typically developing (including 20 children who screened negative on the M-CHAT and STAT and did not participate in a full evaluation). Data were also collected from 38 children selected at random from the norm sample of the BASC-2 Parent Rating Scales Preschool Form to serve as a control group.

Table 1 displays the subject characteristics of children who were included in the current study. The total sample for the study consisted of 99 males and 59 females. The average age at time of evaluation was 36 months (range = 25 - 37 months, $SD = 11$ months). Chi-square and one-way analyses of variance (ANOVAs) were conducted to identify potential differences in subject characteristics amongst the five groups. Chi-square analyses revealed no significant group differences with respect to sex ($\chi^2 = 7.219$, $p = .125$). A one-way ANOVA also indicated that chronological age did not differ by study group ($F (3, 113) = 1.387$, $p = .241$), nor did ethnicity differ significantly by study group ($F (3, 147) = 2.340$, $p = .076$).

**Procedures Utilized for Research**
Caregivers of participants completed the M-CHAT at 18- and 24-month well child visits or early intervention provider visits. M-CHAT forms were sent to the corresponding institutions (UCONN and GSU) and scored by research staff. If M-CHAT results indicated risk for ASDs, a member of the research team called the caregiver to administer the M-CHAT Follow-up Interview (FUI; Robins, Fein, & Barton, 1999b). The purpose of the FUI is to determine if M-CHAT responses were accurate, and to gain additional information regarding potential risk for ASDs. If risk for ASDs was still indicated after the FUI, the family was invited for a free, comprehensive clinical evaluation. The full clinical evaluation includes administration of the Mullen Scales of Early Learning (MSEL; Mullen, 1995), the Vineland Adaptive Behavior Scales-Second Edition (VABS-II; Sparrow et al., 2004), the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2005), the Autism Diagnostic Interview-Revised (ADI-R; Rutter et

Table 1

Descriptive Statistics for Demographic Variables for Total Sample and Diagnostic Categories

<table>
<thead>
<tr>
<th></th>
<th>Autism</th>
<th>PDD-NOS</th>
<th>Other DD</th>
<th>M-CHAT Typicals</th>
<th>Norm Group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>16</td>
<td>27</td>
<td>14</td>
<td>22</td>
<td>20</td>
<td>99</td>
</tr>
<tr>
<td>Females</td>
<td>7</td>
<td>8</td>
<td>14</td>
<td>12</td>
<td>18</td>
<td>59</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>8</td>
<td>20</td>
<td>11</td>
<td>21</td>
<td>28</td>
<td>88</td>
</tr>
<tr>
<td>Black</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>36.74</td>
<td>37.57</td>
<td>32.07</td>
<td>34.21</td>
<td>37.18</td>
<td>35.66</td>
</tr>
<tr>
<td>SD</td>
<td>11.15</td>
<td>11.37</td>
<td>8.70</td>
<td>11.35</td>
<td>11.86</td>
<td>11.09</td>
</tr>
</tbody>
</table>

...
al., 2003), The Toddler-ADI-R (Kim & Lord, 2011), or the Toddler ASD Symptom
Interview (Fein et al., 2010), and the Childhood Autism Rating Scale (CARS; Schopler et
al., 1988). Evaluations were conducted by trained graduate clinicians, and a licensed
psychologist or developmental pediatrician. If the child was at least 24 months old at the
time of the evaluation, the caregiver completed the Behavioral Assessment System for
Children, Second Edition Parent Rating Scales Preschool Form (BASC-2 PRS-P) prior to
the appointment. The examiners then discussed results of the measures administered and
behaviors observed during the evaluation, and clinical judgment was used to determine
diagnosis. If the child was older than 22 months but less than 24 months, at the time of
the evaluation, a member of the research staff mailed the BASC-2 PRS-P to the caregiver
to be completed once the child was at least 24 months of age. BASC-2 scores for three
participants were mailed in after diagnosis.

All families who received a comprehensive evaluation after initial screen positive
results were invited back for a re-evaluation around the time of the child’s fourth birthday
(42 to 60 months). The same measures were administered to determine current levels of
functioning and progress since the child’s initial evaluation. Examiners did not look at the
diagnosis given at the initial evaluation until after determination of the current diagnosis
was made.

As stated previously, a random sample of parents of children who were
determined to be not at risk following administration of M-CHAT follow-up interview
were invited to have their child come in for a secondary screening measure, the Screening
Tool for Autism in Two-Year-Olds (STAT; Stone & Ousley, 1997). If the child was not
at risk on the STAT, parents were given immediate feedback about their results. If results
of the second screening measure indicated risk for ASDs, the family was invited back for the comprehensive evaluation. Parents of children who screened negative on the M-CHAT and the STAT were asked to complete the BASC-2 PRS-P as part of the typically developing sample.

**Instruments**

**Modified Checklist for Autism in Toddlers (M-CHAT).** The Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 1999; Robins et al., 2001) is a 23-item yes/no parent report checklist designed to screen for risk of autism spectrum disorders in children ages 16 to 30 months. Failing any three items, or failing two of the six critical items, is considered a positive screen. The M-CHAT FUI is administered to all participants who screen positive on the M-CHAT, where failed items are re-administered by trained research personnel and follow-up questions are asked to clarify responses and obtain examples of the child's behavior. Reliability estimates of the M-CHAT report adequate internal consistency for both the screener as a whole and the six critical items (alpha=.85; Robins et al. 2001). Positive predictive value of the M-CHAT in a sample of both low-risk children ages 14 to 27 months was estimated to be .57 using both the questionnaire and FUI (Robins, 2008). The Modified Checklist for Autism in Toddlers-Revised (MCHAT-R; Robins, Fein, & Barton, 2009) is a revised version of the M-CHAT. Items from the M-CHAT were reworded on the M-CHAT-R to enhance clarity, and three items from the original questionnaire were dropped due to poor performance. Item order was also changed to avoid response bias. Validity studies of the M-CHAT-R are currently being conducted.
**Autism Diagnostic Observation Schedule (ADOS).** The ADOS (Lord et al. 2005) is a semistructured assessment of communication, play, and social interaction designed to measure the presence of autism spectrum disorders in individual. Activities are designed to elicit a range of communicative and social exchanges. In an attempt to minimize over-diagnosis on the basis of limited expressive language, the authors developed four modules based on differing levels of language acquisition. Module 1 and Module 2 are used most frequently in the assessment of toddlers and preschoolers; Module 1 is designed to be used with children who are pre-verbal or speaking single words, while Module 2 contains activities suitable for children who have acquired phrase speech. The authors of the ADOS report adequate reliability estimates for internal consistency, inter-rater reliability, and test-retest reliability (Lord et al., 2000). Additionally, Lord and Corsello (2005) found that the ADOS-G demonstrated convergent validity with the ADI-R and had good construct validity with DSM-IV criteria for Pervasive Developmental Disorders.

**Screening Tool for Autism in Two-Year-Olds (STAT).** The Screening Tool for Autism in Two-Year-Olds (STAT: Stone & Ousley, 1997) is a brief measure designed to identify the possible presence of autism in children aged 24 to 35 months. Similar to the ADOS, items are administered during a play-like interaction that assess social and communicative behaviors. The STAT contains 12 items and takes approximately 20 minutes to administer. Preliminary data collected by the authors revealed acceptable levels of sensitivity (.83) and specificity (.86). Subsequent research conducted by Stone, Coonrod, Turner, and Pozdol (2004) found the STAT to have good inter-observer agreement ($r=1.00$) and test-retest reliability ($r=.90$). These authors also found high
agreement (.95) between scores on the STAT and ADOS classifications. The STAT was recently validated for use with children under the age of 24 months (Stone, McMahon, & Henderson, 2008).

**Autism Diagnostic Interview-Revised (ADI-R).** The Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994) is a semi-structured diagnostic interview for use with parents or caregivers of people with autism. The ADI-R elicits information related to developmental history as well as current functioning for individuals from early childhood (i.e., with a mental age of at least 24 months) through adulthood. Three domains of development are assessed using the ADI-R: communication, social relatedness, and stereotyped interests and behaviors. The ADI-R contains 93 items, 42 of which comprise the diagnostic algorithm based on the criteria established by the Diagnostic and Statistical Manual of Mental Disorders-4th Edition (DSM-IV; American Psychiatric Association, 1994). Based on scores in the four subdomains of the final diagnostic algorithm (i.e., communication, social relatedness, stereotyped interests and behaviors, and developmental delays observed before child's 3rd birthday), the child is classified as having autism or no autism. Unlike the ADOS-G, the ADI-R does not provide separate cut-off scores for autism spectrum disorders and autistic disorder.

Inter-rater reliability coefficients were between .64 and .89 for ADI-R algorithm items scored on children 36-59 months of age. Tests of discriminant validity showed that the ADI-R algorithm items accurately distinguished between children with autism and those with language impairment or intellectual disability (Lord, Rutter, & Le Couteur, 1994).
**Toddler ASD Symptom Interview.** The Toddler ASD Symptom Interview (Fein et al., 2010) is a comprehensive interview that assesses three domains of development associated with autism spectrum disorders: social reciprocity, communication, and restricted, repetitive, and stereotyped behaviors. The Toddler ASD Symptom Interview is based on DSM-IV criteria for pervasive developmental disorders. Principal investigators of the current study developed the interview as an alternative to the ADI-R, due to evidence of poor reliability and validity of the ADI-R and ADI-R Toddler version when used with children under the age of four (Charwaska et al. 2007; Cox et al. 1999; Gray, Tonge, & Sweeney, 2008).

**Vineland Adaptive Behavior Scales, Second Edition (VABS-II).** The VABS-II Survey Interview Form (Sparrow, Balla, & Cichetti, 2005) is a semi-structured interview administered to parents which assesses skills in the areas of communication, socialization, daily living (i.e., self-help) skills, and motor skills for children ages birth to 18 years. The four domains are further divided into subdomains. Parents are asked to rate how well and/or often the child is able to complete tasks independently using ratings of 0 (never), 1 (sometimes or partially), and 2 (usually). Raw scores are converted to standard scores, percentile ranks, and age equivalents. Internal consistency estimates for parents of 0-5 year old children are .89-.93 for VABS-II domains and .97 for the adaptive behavior composite.

**Childhood Autism Rating Scale (CARS).** The CARS (Schopler, Reichler, & Renner, 1980; 1988) is a standardized measure used to assess symptoms of autism in children ages two years and older. Evaluators incorporate direct observations of behavior along with parent report to rate children on 15 items, including relating to people,
imitation, emotional response, adaptation to change, visual and listening responses, verbal and nonverbal communication, body use, object use, sensory response, fear or nervousness, activity level, intellectual response, and general clinical impressions. Items are scored on a 7-point Likert scale rated from 1 to 4 in half-point increments. Item scores are then summed and classify the child according to severity of autistic symptoms: severe autistic, mildly-moderately autistic, or non-autistic. Reliability estimates for the CARS include an internal consistency of .94, test-retest reliability of .88, and inter-rater reliability of .71. Criterion related validity established by correlating total CARS scores and general clinical ratings of autism severity resulted in a correlation of .84. The most updated version of the CARS, the Childhood Autism Rating Scale, Second Edition-Standard Version (CARS2-ST; Schopler, Van Bourgondien, Wellman, & Love, 2010) is used with the most current participants. The CARS2-ST retained the original content and recommended cutoff values of the CARS.

**Mullen Scales of Early Learning (MSEL).** The MSEL (Mullen, 1995) is a standardized measure of cognitive and motor development in children from birth to 68 months of age, including fine and gross motor skills, receptive and expressive language, and visual reception (i.e., problem-solving). Raw scores obtained on items from each domain are converted to T-scores, percentile ranks, and age equivalents. A composite score is also derived utilizing scores from the cognitive domains (Visual Reception, Fine Motor, Receptive Language, and Expressive Language). Internal consistency reliability estimates for the five MSEL scales range from .75 to .83, and .91 for the composite. Inter-rater reliability correlations ranged from .91 to .99 for raters of age groups between 1 and 44 months. Tests of concurrent validity showed that the MSEL Early Learning
Composite (ELC) was correlated (.70) with Bayley Mental Developmental Index (MDI), and the Gross Motor scale was strongly correlated with the Bayley Psychomotor Development Index (.76). The MSEL Fine Motor scale was strongly correlated with the Peabody Fine Motor Scale, with correlations ranging from .65 to .82 for children aged 6 to 36 months. The language scales of the MSEL demonstrated correlations with the Preschool Language Assessment: the Receptive Language scale had a stronger correlation with Auditory Comprehension (.85) than Verbal Ability (.72), while the Expressive Language scale was more strongly correlated with Verbal Ability (.80) than Auditory Comprehension (.72).

**Behavior Assessment System for Children-Second Edition (BASC-2).** The BASC-2 (Reynolds and Kamphaus, 2004) is a multidimensional assessment system that evaluates both clinical and adaptive aspects of behavior and emotional functioning. Parents, teachers, and students themselves can complete behavior rating scales; however, only the Parent Rating Scale-Preschool Form (PRS-P) is used in this study and will be described here. Each PRS item describes a behavior which parents may observe in their children, and is rated on a four-point frequency scale (i.e., 0=Never, 1=Sometimes, 2=Often, and 3=Almost Always). Item raw scores are summed and converted into standardized T scores with a mean of 50 and a standard deviation of 10 for interpretation. For the clinical scales, higher scores represent more problematic behaviors, with T scores between 60 and 69 considered “at-risk”, and T scores of 70 or above being clinically significant. On the adaptive scales, lower scores are indicative of deficits, with T scores between 31 and 40 falling in the at-risk range and scores equal to or less than 30 considered clinically significant. The BASC-2 PRS-P (ages 2-5) reports T-scores for
behaviors on the following scales: Hyperactivity, Aggression, Anxiety, Depression, Atypicality, Withdrawal, Somatization, Attention Problems, Adaptability, Social Skills, Activities of Daily Living, and Functional Communication. In addition, T-scores are reported for seven content scales: Anger Control, Bullying, Developmental Social Disorders, Emotional Self-Control, Executive Functioning, Negative Emotionality, and Resiliency.

Authors of the BASC-2 PRS-P report adequate reliability and validity. Individual scales of the BASC-2 PRS-P have a median test-retest reliability of 0.77, and median inter-rater reliability of .74. The Behavioral Symptoms Index of the BASC-2 PRS-P is also highly correlated with the Achenbach System of Empirically Based Assessment Child Behavior Checklist for Ages 1.5-5 Total Problems score ($r = 0.75$).

Data Analyses

Data analyses addressed the following research questions:

1. Do children diagnosed with different developmental delays (AD, PDD-NOS, Other DDs) obtain significantly different $T$-scores on the clinical scales, adaptive scales, and Developmental Social Disorders content scale of the BASC-2 PRS-P in comparison with typically developing children?

2. (A) Does the BASC-2 DSD scale demonstrate adequate sensitivity and specificity for classifying toddlers and preschoolers with ASDs? (B) Which cut-off score on the BASC-2 DSD scale most accurately classifies young children with ASD?

3. Does the BASC-2 DSD scale demonstrate concurrent validity with total scores of the ADOS (Communication + Reciprocal Social Interaction) and the total score CARS?
Research Question 1. To rule out pre-existing subject characteristic differences as potential explanations for results, between-group comparisons were conducted for age and sex. Chi-square and one-way analyses of variance (ANOVA) were conducted to identify potential differences in subject characteristics amongst the five groups (Autistic Disorder, PDD-NOS, Other Developmental Delays, M-CHAT Typical, and Norm Group Typical). Next, Pearson correlations were conducted to examine the associations between age and sex and T-scores for the BASC-2 clinical scales, adaptive scales, and Developmental Social Disorders content scale. These analyses were collapsed across all diagnostic groups.

*F*-test comparisons of BASC-2 PRS T-score were conducted between pairs of diagnostic groups to determine if certain diagnostic groups could be combined in further analyses. *F*-test comparisons were first conducted to compare participants diagnosed with an ASD (i.e., Autistic Disorder or PDD-NOS). Analyses were then conducted between participants diagnosed with Other Developmental Delays (i.e., Global Developmental Delay and Developmental Language Delay). *F*-test comparisons were also conducted to determine if any differences existed between participants in the M-CHAT Typical group, which included subjects labeled as having “no diagnosis”, those diagnosed as typically developing, and those who screened negative on the STAT but did not participate in the full evaluation. Finally, *F*-test comparisons were conducted to determine if significant differences were observed on BASC-2 T-scores between participants from the M-CHAT Typical group and those randomly selected from the BASC-2 norming sample.

*F*-test comparisons of BASC-2 PRS T-scores were conducted across groups using MANCOVA to control for subject characteristics that might act as alternative...
explanations for differences. Corrections for multiple comparisons were conducted using the Bonferroni technique. The first set of analyses compared $T$-scores on the BASC-2 clinical scales across four groups: participants diagnosed with an ASD, those diagnosed with other developmental delays, typically developing children from the M-CHAT study sample, and typically developing subjects from the BASC-2 norming sample. $F$-test comparisons were then run between these for groups for the adaptive scales and Developmental Social Disorders (DSD) scale of the BASC-2. In the second set of analyses, comparisons on the clinical, adaptive, and DSD scales were conducted across five groups, separating the ASD group into participants diagnosed with Autistic Disorder and participants diagnosed with PDD-NOS.

**Research Question 2.** To answer research question 2, screening sensitivity and specificity estimates were calculated based on the current sample for the Developmental Social Disorders Scale. The predictive power of the DSD scale was estimated using the Positive Predictive Value (PPV) and Negative Predictive Value (NPV). PPV indicates the proportion of children with scores above the clinical cut-off who are diagnosed with an ASD (true positives). NPV indicates the proportion of clinical cases with negative screens that are not diagnosed with an ASD (true negatives). Results for three different screening cut scores (i.e., $T$ scores of 60, 65, and 70) were reported.

**Research Question 3.** Pearson product moment correlations were conducted to identify associations between the BASC-2 Developmental Social Disorders Scale $T$-Score and the total scores (Communication + Reciprocal Social Interaction) of the ADOS and the total score CARS. Values of $\pm .1$ represent a small effect, $\pm .3$ represent a medium effect, and values $\geq \pm .5$ represent a large effect (Field, 2009).
CHAPTER III: RESULTS

Correlations Between Participant Characteristics and BASC-2

Pearson correlations were conducted to examine the associations between age and sex and the BASC-2 clinical (see Table 2) and adaptive scale T-scores (see Table 3) collapsed across all diagnostic groups. There was a significant positive correlation between age and five BASC-2 scales: Hyperactivity ($r = .250, p = .002$), Aggression ($r = .253, p = .001$), Anxiety ($r = .195, p = 0.015$), Social Skills ($r = .201, p = .011$), and Functional Communication ($r = .227, p = .004$). Sex differences were not significant for most BASC-2 scales; however, males obtained higher T-scores on the Depression scale than females ($r = .186, p = .020$). Additionally, although not statistically significant, there was a notable trend towards males obtaining higher T-scores on the DSD scale.

Table 2

Correlation Matrix of Demographic Variables by BASC-2 Clinical Scale Scores for Total Sample (n=158)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hyp</th>
<th>Agg</th>
<th>Anx</th>
<th>Dep</th>
<th>Som</th>
<th>Atp</th>
<th>Wdl</th>
<th>Atn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>Pearson $r$</td>
<td>.250</td>
<td>.253</td>
<td>.195</td>
<td>.005</td>
<td>.078</td>
<td>.006</td>
<td>-.024</td>
</tr>
<tr>
<td></td>
<td>$p$-value</td>
<td>.002</td>
<td>.001</td>
<td>.015</td>
<td>.948</td>
<td>.331</td>
<td>.940</td>
<td>.766</td>
</tr>
<tr>
<td>Sex</td>
<td>Spearman’s rho</td>
<td>.096</td>
<td>.112</td>
<td>.014</td>
<td>.186</td>
<td>.092</td>
<td>.008</td>
<td>-.029</td>
</tr>
<tr>
<td></td>
<td>$p$-value</td>
<td>.229</td>
<td>.162</td>
<td>.867</td>
<td>.020</td>
<td>.249</td>
<td>.917</td>
<td>.720</td>
</tr>
</tbody>
</table>

Note: $n =$ number of participants in total sample; Hyp = Hyperactivity; Agg = Aggression; Anx = Anxiety; Dep = Depression; Som = Somatization; Atp = Atypicality; Wdl = Withdrawal; Atn = Attention Problems
Table 3

Correlation Matrix of Demographic Variables by BASC-2 Adaptive Scale Scores and DSD Scale Score for Total Sample (n=158)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adt</th>
<th>Skl</th>
<th>Adl</th>
<th>Fun</th>
<th>DSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>Pearson r</td>
<td>-.068</td>
<td>.201</td>
<td>.074</td>
<td>.227</td>
</tr>
<tr>
<td></td>
<td>p-value</td>
<td>.399</td>
<td>.011</td>
<td>.357</td>
<td>.004</td>
</tr>
<tr>
<td>Sex</td>
<td>Spearman’s rho</td>
<td>-.025</td>
<td>-.072</td>
<td>.007</td>
<td>-.037</td>
</tr>
<tr>
<td></td>
<td>p-value</td>
<td>.754</td>
<td>.371</td>
<td>.931</td>
<td>.647</td>
</tr>
</tbody>
</table>

Note: n = number of participants in total sample; Adt = Adaptability; Skl = Social Skills, Adl = Activities of Daily Living; Fun = Functional Communication; DSD = Developmental Social Disorders

Differences Within Study Groups

Autistic Disorder vs. PDD-NOS

First, individual, independent samples t-tests were run for each scale to retain maximum statistical power and reduce the possibility of making a Type I error. Results of the individual t-tests revealed no significant differences between groups. Next, MANCOVA were conducted to determine if significant differences clinical scale T-scores on the clinical scales were observed between the two groups when age was controlled. Corrections for multiple comparisons were made using the Bonferroni technique. Table 4 displays the means and standard deviations of the BASC-2 clinical scale scores for participants diagnosed with Autistic Disorder and those diagnosed with PDD-NOS Analysis revealed that the overall multivariate effect between diagnostic classifications. Table 4 displays the means and standard deviations of the BASC-2 clinical scale scores for participants diagnosed with Autistic Disorder and those diagnosed with PDD-NOS Analysis revealed that the overall multivariate effect between diagnostic
Table 4

*MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (Autism and PD-NOS groups)*

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>Autism (n=23) Mean (SD)</th>
<th>PDD-NOS (n=35) Mean (SD)</th>
<th>F value</th>
<th>p value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>55.13 (12.21)</td>
<td>58.08 (13.93)</td>
<td>.438</td>
<td>.511</td>
<td>.008</td>
</tr>
<tr>
<td>Aggression</td>
<td>48.00 (10.98)</td>
<td>51.67 (13.57)</td>
<td>.896</td>
<td>.348</td>
<td>.016</td>
</tr>
<tr>
<td>Anxiety</td>
<td>44.91 (10.05)</td>
<td>46.78 (14.07)</td>
<td>.272</td>
<td>.604</td>
<td>.005</td>
</tr>
<tr>
<td>Depression</td>
<td>52.70 (14.78)</td>
<td>53.86 (11.38)</td>
<td>.092</td>
<td>.763</td>
<td>.002</td>
</tr>
<tr>
<td>Somatization</td>
<td>49.35 (9.50)</td>
<td>50.42 (10.75)</td>
<td>.076</td>
<td>.784</td>
<td>.001</td>
</tr>
<tr>
<td>Atypicality</td>
<td>66.65 (14.04)</td>
<td>64.67 (16.39)</td>
<td>.353</td>
<td>.555</td>
<td>.006</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>59.83 (11.87)</td>
<td>57.50 (10.55)</td>
<td>.545</td>
<td>.463</td>
<td>.010</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>63.78 (10.04)</td>
<td>60.78 (10.04)</td>
<td>1.731</td>
<td>.194</td>
<td>.030</td>
</tr>
</tbody>
</table>

classification groups was not significant \([Pillai’s Trace = .107, F (8, 49) = .732, and p = .663]\).

Independent samples t-tests run for the adaptive and Developmental Social Disorders (DSD) scales revealed that there was a significant difference between the Autism group and the PDD-NOS group on the Social Skills scale, with children in the Autism group obtaining significantly lower scores on the Social Skills scale than children diagnosed with PDD-NOS \((t = -2.704, p = .009)\). MANCOVA were also conducted to determine if significant differences between adaptive scale T-scores and the DSD scale were observed between the two groups when age was controlled (see Table 5). While analysis did not reveal an overall significant multivariate effect between diagnostic classification groups \([Pillai’s Trace = .116, F (5, 52) = 1.369, and p = .251]\), corrections made for multiple comparisons using the Bonferroni technique again revealed a
Table 5

MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (Autism and PD-NOS groups)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>Autism (n=23) Mean (SD)</th>
<th>PDD-NOS (n=35) Mean (SD)</th>
<th>F-value</th>
<th>p-value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>40.91 (9.27)</td>
<td>43.42 (11.65)</td>
<td>.825</td>
<td>.368</td>
<td>.015</td>
</tr>
<tr>
<td>Social Skills</td>
<td>32.52 (6.19)</td>
<td>38.11 (9.69)</td>
<td>5.674</td>
<td>.021</td>
<td>.092</td>
</tr>
<tr>
<td>Activities of Daily Living</td>
<td>36.74 (9.87)</td>
<td>38.81 (12.42)</td>
<td>.620</td>
<td>.434</td>
<td>.011</td>
</tr>
<tr>
<td>Functional Communication</td>
<td>34.65 (1.20)</td>
<td>37.11 (10.14)</td>
<td>1.061</td>
<td>.307</td>
<td>.019</td>
</tr>
<tr>
<td>Developmental Social Disorders</td>
<td>69.74 (7.86)</td>
<td>65.39 (11.43)</td>
<td>2.931</td>
<td>.092</td>
<td>.050</td>
</tr>
</tbody>
</table>

significant difference in T-scores obtained between participants diagnosed with Autism and those diagnosed with PDD-NOS on the Social Skills scale ($F (1, 56) = 5.674, p = .021$).

Global Developmental Delay vs. Developmental Language Delay

Table 6 displays the means and standard deviations of the BASC-2 clinical scale scores for participants diagnosed with Global Developmental Delay and those diagnosed with Developmental Language Delay. Again, individual independent samples t-tests were conducted for each scale to retain maximum statistical power and reduce the chances of making a Type I error. Independent samples t-tests revealed no significant differences between groups. MANCOVA were then conducted to determine if significant differences clinical scale T-scores on the clinical scales were observed between the two groups when
Table 6

MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (GDD and Lang Delay groups)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>GDD (n=20) Mean (SD)</th>
<th>Lang Delay (n=8) Mean (SD)</th>
<th>F value</th>
<th>p value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>49.75 (11.89)</td>
<td>50.50 (7.56)</td>
<td>.091</td>
<td>.766</td>
<td>.004</td>
</tr>
<tr>
<td>Aggression</td>
<td>45.35 (9.70)</td>
<td>46.75 (7.74)</td>
<td>.361</td>
<td>.554</td>
<td>.015</td>
</tr>
<tr>
<td>Anxiety</td>
<td>43.10 (7.09)</td>
<td>46.00 (7.92)</td>
<td>.633</td>
<td>.434</td>
<td>.026</td>
</tr>
<tr>
<td>Depression</td>
<td>52.00 (11.96)</td>
<td>51.13 (13.40)</td>
<td>.040</td>
<td>.843</td>
<td>.002</td>
</tr>
<tr>
<td>Somatization</td>
<td>47.95 (12.17)</td>
<td>49.00 (6.11)</td>
<td>.039</td>
<td>.845</td>
<td>.002</td>
</tr>
<tr>
<td>Atypicality</td>
<td>59.95 (12.38)</td>
<td>55.13 (11.45)</td>
<td>.148</td>
<td>.704</td>
<td>.006</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>55.45 (9.52)</td>
<td>54.50 (11.25)</td>
<td>.010</td>
<td>.921</td>
<td>.000</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>57.50 (10.43)</td>
<td>57.25 (9.27)</td>
<td>.082</td>
<td>.777</td>
<td>.003</td>
</tr>
</tbody>
</table>

age was controlled. Corrections for multiple comparisons were made using the Bonferroni technique. Analysis revealed that the overall multivariate effect between diagnostic classification groups was not significant [Pillai’s Trace = .112, F (8,17) = .269, and \( p = .968 \)].

Independent samples \( t \)-tests and were also conducted to determine if significant differences existed between adaptive scale T-scores and the DSD scale, and again revealed no significant differences between groups. A MANCOVA was then conducted to determine if significant differences were observed between the two groups when age was controlled (see Table 7). Corrections for multiple comparisons were made using the Bonferroni technique. Again, analysis revealed that the overall multivariate effect between diagnostic classification groups was not significant [Pillai’s Trace = .205, \( F (5, 21) = 1.086, \) and \( p = .397 \)].
Table 7

MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (GDD and Lang Delay groups)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>GDD (n=20) Mean (SD)</th>
<th>Lang Delay (n=8) Mean (SD)</th>
<th>F-value</th>
<th>p-value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>43.85 (10.71)</td>
<td>45.63 (16.53)</td>
<td>.133</td>
<td>.719</td>
<td>.005</td>
</tr>
<tr>
<td>Social Skills</td>
<td>32.85 (5.76)</td>
<td>38.38 (8.86)</td>
<td>3.258</td>
<td>.083</td>
<td>.115</td>
</tr>
<tr>
<td>Activities of Daily Living</td>
<td>40.65 (10.75)</td>
<td>49.13 (13.65)</td>
<td>2.609</td>
<td>.119</td>
<td>.095</td>
</tr>
<tr>
<td>Functional Communication</td>
<td>36.05 (4.10)</td>
<td>37.38 (6.59)</td>
<td>.062</td>
<td>.805</td>
<td>.002</td>
</tr>
<tr>
<td>Developmental Social Disorders</td>
<td>61.70 (6.30)</td>
<td>59.63 (8.37)</td>
<td>.208</td>
<td>.652</td>
<td>.008</td>
</tr>
</tbody>
</table>

M-CHAT Typical Group

The M-CHAT study group included children labeled as having “no diagnosis,” those diagnosed as “typically developing,” and children who did not participate in the full diagnostic evaluation, but screened negative on the STAT. First, individual ANOVAs were run comparing these three groups to retain maximum statistical power given the small number of subjects in each group. Separate one-way ANOVAs conducted for the clinical scales revealed no significant differences between participants. MANCOVA (presented in Table 8) conducted for the BASC-2 clinical scales revealed no significant differences between participants classified as typically developing, those classified as having no diagnosis, and those who screened negative on the STAT when age was controlled [Pillai’s Trace = .588, F (16, 46) = 1.197, p = .306].
Table 8

MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (M-CHAT Typical groups)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>No Dx (n=9) Mean (SD)</th>
<th>STAT Neg (n=20) Mean (SD)</th>
<th>Typical (n=5) Mean (SD)</th>
<th>F value</th>
<th>p-value</th>
<th>Partial Eta^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>54.89 (12.43)</td>
<td>48.70 (5.25)</td>
<td>48.00 (5.81)</td>
<td>2.468</td>
<td>.102</td>
<td>.145</td>
</tr>
<tr>
<td>Aggression</td>
<td>49.67 (7.89)</td>
<td>47.05 (6.46)</td>
<td>51.20 (8.14)</td>
<td>.470</td>
<td>.630</td>
<td>.031</td>
</tr>
<tr>
<td>Anxiety</td>
<td>53.33 (9.00)</td>
<td>49.40 (9.58)</td>
<td>54.25 (12.61)</td>
<td>.796</td>
<td>.461</td>
<td>.052</td>
</tr>
<tr>
<td>Depression</td>
<td>53.00 (7.19)</td>
<td>50.40 (9.89)</td>
<td>44.25 (9.32)</td>
<td>3.059</td>
<td>.062</td>
<td>.174</td>
</tr>
<tr>
<td>Somatization</td>
<td>50.78 (5.43)</td>
<td>49.70 (9.42)</td>
<td>48.00 (8.72)</td>
<td>.564</td>
<td>.575</td>
<td>.037</td>
</tr>
<tr>
<td>Atypicality</td>
<td>53.00 (11.98)</td>
<td>56.10 (9.14)</td>
<td>47.80 (6.65)</td>
<td>2.990</td>
<td>.066</td>
<td>.171</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>55.11 (14.68)</td>
<td>50.10 (6.07)</td>
<td>52.60 (12.22)</td>
<td>.263</td>
<td>.770</td>
<td>.018</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>55.44 (7.54)</td>
<td>51.00 (7.31)</td>
<td>51.00 (8.37)</td>
<td>.693</td>
<td>.508</td>
<td>.046</td>
</tr>
</tbody>
</table>
Separate one-way ANOVAs conducted for the adaptive and DSD scales revealed no significant differences between groups. MANCOVA for the BASC-2 adaptive scales and DSD scale (presented in Table 9) also revealed no significant differences between the three groups when age was controlled \([\text{Pillai’s Trace} = .368, F (10, 54) = 1.219, p = .300]\). Estimates of effect size were also small; therefore, it was determined that these groups could be combined in further analyses and were labeled as the “M-CHAT Typical” group.

**M-CHAT Typical vs. Norm Sample Group**

Independent samples t-tests were then conducted to determine if there were significant differences on BASC-2 scale scores between participants in the M-CHAT Typical group and subjects selected from the BASC-2 norming sample population. Significant differences were observed between participants in the M-CHAT Typical group and participants from the norming sample for the Anxiety \((t = 2.038, p = .045)\), Atypicality \((t = 2.444, p = .017)\), Withdrawal \((t = 2.287, p = .025)\), and DSD \((t = 4.428, p < .001)\) scales. MANCOVA were also conducted to determine if significant differences between clinical scales T-scores were observed between typically developing participants from the M-CHAT study sample and children from the BASC-2 norming sample when age was controlled (see Table 10). Analysis revealed an overall significant multivariate effect between groups \([\text{Pillai’s Trace} = .216, F (8, 61) = 2.097, p = .050]\). Post-hoc analysis using the Bonferroni technique revealed that participants from the M-CHAT study sample obtained significantly higher T-scores than those from the norming group.
Table 9

**MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (M-CHAT Typical groups)**

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>No Dx (n=9) Mean (SD)</th>
<th>STAT Neg (n=20) Mean (SD)</th>
<th>Typical (n=5) Mean (SD)</th>
<th>F value</th>
<th>p-value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>50.56 (6.75)</td>
<td>50.10 (6.88)</td>
<td>52.00 (10.32)</td>
<td>1.276</td>
<td>.294</td>
<td>.078</td>
</tr>
<tr>
<td>Social Skills</td>
<td>49.22 (11.58)</td>
<td>47.80 (7.53)</td>
<td>44.80 (3.56)</td>
<td>1.308</td>
<td>.285</td>
<td>.080</td>
</tr>
<tr>
<td>Activities of Daily Living</td>
<td>44.11 (14.23)</td>
<td>45.45 (12.25)</td>
<td>45.40 (10.36)</td>
<td>.242</td>
<td>.787</td>
<td>.016</td>
</tr>
<tr>
<td>Functional Communication</td>
<td>48.78 (8.29)</td>
<td>44.50 (7.07)</td>
<td>45.00 (8.22)</td>
<td>1.028</td>
<td>.370</td>
<td>.064</td>
</tr>
<tr>
<td>Developmental Social Disorders</td>
<td>55.44 (9.90)</td>
<td>53.85 (5.67)</td>
<td>57.80(8.87)</td>
<td>.077</td>
<td>.926</td>
<td>.005</td>
</tr>
</tbody>
</table>
Table 10

MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (M-CHAT Typical and Norm Group)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>M-CHAT (n=34) Mean (SD)</th>
<th>Norm Group (n=38) Mean (SD)</th>
<th>F value</th>
<th>p value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>50.24 (8.22)</td>
<td>49.61 (9.70)</td>
<td>.352</td>
<td>.555</td>
<td>.005</td>
</tr>
<tr>
<td>Aggression</td>
<td>48.52 (7.11)</td>
<td>50.37 (10.42)</td>
<td>.333</td>
<td>.566</td>
<td>.005</td>
</tr>
<tr>
<td>Anxiety</td>
<td>51.06 (9.70)</td>
<td>46.50 (9.14)</td>
<td>7.019</td>
<td>.010</td>
<td>.094</td>
</tr>
<tr>
<td>Depression</td>
<td>50.36 (9.26)</td>
<td>48.97 (10.62)</td>
<td>.641</td>
<td>.426</td>
<td>.009</td>
</tr>
<tr>
<td>Somatization</td>
<td>49.94 (8.30)</td>
<td>45.66 (10.71)</td>
<td>3.760</td>
<td>.057</td>
<td>.052</td>
</tr>
<tr>
<td>Atypicality</td>
<td>54.48 (9.68)</td>
<td>49.18 (6.97)</td>
<td>7.350</td>
<td>.008</td>
<td>.098</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>51.88 (9.96)</td>
<td>47.13 (7.43)</td>
<td>5.793</td>
<td>.019</td>
<td>.079</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>52.33 (7.61)</td>
<td>49.00 (7.84)</td>
<td>3.670</td>
<td>.060</td>
<td>.051</td>
</tr>
</tbody>
</table>

on three clinical scales: Anxiety ($F(1,68) = 7.019, p = .010$), Atypicality ($F(1,68) = 7.350, p = .008$), and Withdrawal ($F(1,68) = 5.793, p = .019$).

Multivariate analyses of covariance (presented in Table 11) also revealed statistically significant differences on the Developmental Social Scale between typically developing participants from the M-CHAT study sample and typically developing children from the BASC-2 norming sample when age was controlled ($F(1,69) = 20.002, p < .001$). Therefore, these two groups were analyzed separately in further analyses.

**Multivariate Analyses of Covariance (ASDs Collapsed)**

Table 12 displays the means and standard deviations of the BASC-2 clinical scale scores for the ASD, Other Developmental Delays, M-CHAT Typical, and Norm Sample Typical groups. Multivariate analyses of covariance were conducted to determine
### Table 11

**MANCOVA for BASC-2 Adaptive Scales and DSD Scale Using Age (months) as Covariate (M-CHAT Typical and Norm Group)**

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>M-CHAT (n=34) Mean (SD)</th>
<th>Norm Group (n=38) Mean (SD)</th>
<th>$F$ value</th>
<th>$p$-value</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>50.50 (7.19)</td>
<td>50.47 (10.06)</td>
<td>.008</td>
<td>.930</td>
<td>.000</td>
</tr>
<tr>
<td>Social Skills</td>
<td>47.74 (8.28)</td>
<td>47.42 (9.18)</td>
<td>.147</td>
<td>.702</td>
<td>.002</td>
</tr>
<tr>
<td>Activities of Daily Living</td>
<td>45.09 (12.20)</td>
<td>45.55 (9.99)</td>
<td>.000</td>
<td>.998</td>
<td>.000</td>
</tr>
<tr>
<td>Functional Communication</td>
<td>45.71 (7.56)</td>
<td>47.05 (8.51)</td>
<td>.175</td>
<td>.677</td>
<td>.003</td>
</tr>
<tr>
<td>Developmental Social Disorders</td>
<td>54.85 (7.34)</td>
<td>47.13 (7.43)</td>
<td>20.002</td>
<td>&lt;.001</td>
<td>.225</td>
</tr>
</tbody>
</table>

Whether the BASC-2 clinical scales were able to distinguish among diagnostic classification groups when age was controlled. Analysis revealed a significant multivariate effect between diagnostic classification groups [$Pillai’s Trace = .524$, $F (24, 438) = 3.86$, and $p < 0.001$]. For the BASC-2 clinical scales, there was a significant group effect for Hyperactivity [$F (3, 151) = 3.81$, $p = .011$], Anxiety [$F (3, 151) = 2.892$, $p = .037$], Atypicality [$F (3, 151) = 14.95$, $p < .001$], Withdrawal [$F (3, 151) = 10.39$, $p < .001$], Attention Problems [$F (3, 151) = 19.14$, $p < .001$].

Post-hoc analyses using the Bonferroni technique revealed significant differences between the ASD group and the Norm group for Hyperactivity ($p = .016$); however, once corrections for multiple comparisons were made differences between groups were not significant on the Anxiety scale. On the Atypicality scale, participants in the ASD group obtained significantly higher T-scores than participants in the M-CHAT Typical ($p =$...
Table 12

**MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (ASDs Collapsed)**

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>ASD</th>
<th>Other DD</th>
<th>MCHAT Typical</th>
<th>Norm Group</th>
<th>F value</th>
<th>p value</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>56.29 (12.42)</td>
<td>49.96 (10.90)</td>
<td>50.24 (8.22)</td>
<td>49.61 (9.70)</td>
<td>3.811</td>
<td>.011</td>
<td>.070</td>
</tr>
<tr>
<td>Aggression</td>
<td>49.38 (10.90)</td>
<td>45.93 (9.19)</td>
<td>48.52 (7.11)</td>
<td>50.37 (10.42)</td>
<td>.636</td>
<td>.593</td>
<td>.012</td>
</tr>
<tr>
<td>Anxiety</td>
<td>46.05 (12.71)</td>
<td>43.85 (7.27)</td>
<td>51.06 (9.70)</td>
<td>46.50 (9.14)</td>
<td>2.892</td>
<td>.037</td>
<td>.054</td>
</tr>
<tr>
<td>Depression</td>
<td>52.95 (12.31)</td>
<td>51.96 (12.32)</td>
<td>50.36 (9.26)</td>
<td>48.97 (10.62)</td>
<td>1.043</td>
<td>.376</td>
<td>.020</td>
</tr>
<tr>
<td>Somatization</td>
<td>49.52 (9.60)</td>
<td>48.22 (10.82)</td>
<td>49.94 (8.30)</td>
<td>45.66 (10.71)</td>
<td>1.594</td>
<td>.193</td>
<td>.031</td>
</tr>
<tr>
<td>Atypicality</td>
<td>64.62 (14.20)</td>
<td>59.19 (11.89)</td>
<td>54.48 (9.77)</td>
<td>49.18 (6.97)</td>
<td>14.951</td>
<td>&lt;.001</td>
<td>.229</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>58.36 (11.13)</td>
<td>55.22 (10.02)</td>
<td>51.88 (9.96)</td>
<td>47.13 (7.43)</td>
<td>10.393</td>
<td>&lt;.001</td>
<td>.171</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>61.69 (8.84)</td>
<td>57.70 (10.02)</td>
<td>52.33 (7.61)</td>
<td>49.00 (7.84)</td>
<td>19.136</td>
<td>&lt;.001</td>
<td>.275</td>
</tr>
</tbody>
</table>

.001) and Norm groups ($p < 0.001$); however, there were no significant differences between participants in the ASD group and those with Other Developmental Delays on the Atypicality scale ($p = .296$). Participants in the Other Developmental Delay group also obtained significantly higher T-scores on the Atypicality scale than subjects from the Norm Group ($p = .005$). On the Withdrawal scale, subjects in the ASD group obtained significantly higher T-scores than participants from the M-CHAT Typical ($p = .018$) and Norm Group ($p < 0.001$). Participants from the Other Developmental Delay group also obtained higher T-scores on the Withdrawal scale than subjects from the Norm Group.
Participants in the ASD group obtained significantly higher T-scores on the Attention Problems scale than subjects from the M-CHAT Typical and Norm Group \( (p < 0.001) \). Subjects from the Other Developmental Delay group also demonstrated significantly higher T-scores than the Norm Group on the Attention Problems scale \( (p < .001) \).

Table 13 displays the means and standard deviations of the BASC-2 adaptive scale scores for the diagnostic classification groups. MANCOVA were conducted to determine whether the BASC-2 adaptive scales were able to distinguish among diagnostic classification groups when age was controlled. Analysis revealed a significant multivariate effect between diagnostic classification groups \( [Pillai's \, Trace = .442, F (12, 456) = 6.57, \, and \, p < 0.001] \). For the BASC-2 adaptive scales, there was a significant group effect for all four Adaptive scales: Adaptability \( [F (3, 153) = 6.79, \, p < .001] \), Social Skills \( [F (3, 153) = 27.19, \, p < .001] \), Activities of Daily Living \( [F (3, 153) = 4.42, \, p = 0.01] \), and Functional Communication \( [F (3, 153) = 22.96, \, p < .001] \).

Post-hoc analyses using the Bonferroni technique revealed significant differences between participants in the ASD group and subjects in the M-CHAT Typical \( (p = .004) \) and Norm Group \( (p = .002) \) on the Adaptability scale, with the participants from the ASD group obtaining significantly lower T-scores than either of the typically developing groups. On the Social Skills scale, significant differences were obtained between participants in the ASD group and subjects in the M-CHAT Typical \( (p < 0.001) \) and Norm Group \( (p < 0.001) \); however, there were no significant differences between participants in the ASD group and participants diagnosed with Other Developmental Delay. On the Activities of Daily Living scale, subjects in the ASD group scored
Table 13
MANCOVA for BASC-2 Adaptive Scales Using Age (months) as Covariate (ASDs Collapsed)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>ASD</th>
<th>Other DD</th>
<th>MCHAT Typical</th>
<th>Norm Group</th>
<th>F value</th>
<th>p-value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>42.59(10.81)</td>
<td>44.36(12.34)</td>
<td>50.50(7.19)</td>
<td>50.47(10.06)</td>
<td>6.789</td>
<td>&lt;.001</td>
<td>.117</td>
</tr>
<tr>
<td>Social Skills</td>
<td>36.05(8.90)</td>
<td>34.43(7.08)</td>
<td>47.74(8.28)</td>
<td>47.42(9.18)</td>
<td>27.192</td>
<td>&lt;.001</td>
<td>.348</td>
</tr>
<tr>
<td>Activities of Daily Living</td>
<td>38.26(11.37)</td>
<td>43.07(12.03)</td>
<td>45.09(12.20)</td>
<td>45.55(9.99)</td>
<td>4.420</td>
<td>.005</td>
<td>.080</td>
</tr>
<tr>
<td>Functional Communication</td>
<td>36.50(8.37)</td>
<td>36.43(4.97)</td>
<td>45.71(7.56)</td>
<td>47.05(8.51)</td>
<td>22.964</td>
<td>&lt;.001</td>
<td>.310</td>
</tr>
</tbody>
</table>

significantly lower than subjects from the M-CHAT Typical group \( p = .026 \) and the Norm Group \( p = .015 \). Participants from the ASD and Other Developmental Delay groups scored significantly lower on the Functional Communication scale than participants from the M-CHAT Typical and Norm Groups \( p < 0.001 \).

**Multivariate Analyses of Covariance (Autism and PDD-NOS Separated)**

Table 14 displays the means and standard deviations of the BASC-2 clinical scale scores for the Autism, PDD-NOS, Other Developmental Delays, M-CHAT Typical, and Norm Sample Typical groups. Multiple analyses of covariance were conducted to determine whether the BASC-2 clinical scales were able to distinguish among diagnostic classification groups when age was controlled. Analysis revealed a significant multivariate effect between diagnostic classification groups \( [\text{Pillai's Trace} = .561, F (32, 584) = 2.97, p < .001] \). For the BASC-2 clinical scales, there was a significant group effect for Hyperactivity \( [F (4, 150) = 2.94, p = .022] \), Atypicality \( [F (4, 150) = \)
Table 14

**MANCOVA for BASC-2 Clinical Scales Using Age (months) as Covariate (ASDs Separated)**

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>Autism</th>
<th>PDD-NOS</th>
<th>Other DD</th>
<th>M-CHAT Typical</th>
<th>Norm Group</th>
<th>F</th>
<th>p-value</th>
<th>Partial Eta^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>55.13 (12.21)</td>
<td>57.06 (12.67)</td>
<td>49.96 (10.90)</td>
<td>50.24 (8.22)</td>
<td>49.61 (9.70)</td>
<td>2.942</td>
<td>.022</td>
<td>.073</td>
</tr>
<tr>
<td>Aggression</td>
<td>48.00 (10.98)</td>
<td>50.29 (10.91)</td>
<td>45.93 (9.19)</td>
<td>48.52 (7.11)</td>
<td>50.37 (10.42)</td>
<td>.645</td>
<td>.632</td>
<td>.017</td>
</tr>
<tr>
<td>Anxiety</td>
<td>44.91 (10.05)</td>
<td>46.80 (14.28)</td>
<td>43.85 (7.27)</td>
<td>51.06 (9.70)</td>
<td>46.50 (9.14)</td>
<td>2.257</td>
<td>.066</td>
<td>.057</td>
</tr>
<tr>
<td>Depression</td>
<td>52.70 (14.78)</td>
<td>53.11 (10.61)</td>
<td>51.96 (12.32)</td>
<td>50.36 (9.26)</td>
<td>48.97 (10.62)</td>
<td>.782</td>
<td>.539</td>
<td>.020</td>
</tr>
<tr>
<td>Somatization</td>
<td>49.35 (9.50)</td>
<td>49.63 (9.79)</td>
<td>48.22 (10.82)</td>
<td>49.94 (8.30)</td>
<td>45.66 (10.71)</td>
<td>1.189</td>
<td>.318</td>
<td>.031</td>
</tr>
<tr>
<td>Atypicality</td>
<td>66.65 (14.04)</td>
<td>63.29 (14.34)</td>
<td>59.19 (11.89)</td>
<td>54.48 (9.77)</td>
<td>49.18 (6.97)</td>
<td>11.528</td>
<td>&lt;.001</td>
<td>.235</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>59.83 (11.87)</td>
<td>57.40 (10.69)</td>
<td>55.22 (10.02)</td>
<td>51.88 (9.96)</td>
<td>47.13 (7.43)</td>
<td>7.989</td>
<td>&lt;.001</td>
<td>.176</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>63.78 (6.84)</td>
<td>60.31 (9.79)</td>
<td>57.70 (10.02)</td>
<td>52.33 (7.61)</td>
<td>49.00 (7.84)</td>
<td>15.050</td>
<td>&lt;.001</td>
<td>.286</td>
</tr>
</tbody>
</table>
Post-hoc analyses using the Bonferroni technique revealed significant differences between the PDD-NOS group and the Norm group for Hyperactivity ($p = .032$). On the Atypicality scale, participants in the Autism group obtained significantly higher T-scores than participants in the M-CHAT Typical ($p = .002$) and Norm groups ($p < .001$). Participants in the PDD-NOS group also obtained higher T-scores on the Atypicality scale than subjects in the M-CHAT Typical ($p = .022$) and Norm Group ($p < .001$). Participants in the Other Developmental Delay group also obtained significantly higher T-scores on the Atypicality scale than subjects from the Norm Group ($p = .008$). On the Withdrawal scale, subjects in the Autism group obtained significantly higher T-scores than participants from the M-CHAT Typical ($p = .036$) and Norm Group ($p < .001$). Participants from the PDD-NOS group ($p < .001$) and the Other Developmental Delay group ($p = .019$) also obtained higher T-scores on the Withdrawal scale than subjects from the Norm Group. Participants in the Autism group and the PDD-NOS group obtained significantly higher T-scores on the Attention Problems scale than subjects from the M-CHAT Typical and Norm Group ($p < .001$). Subjects from the Other Developmental Delay group also demonstrated significantly higher T-scores than the Norm Group on the Attention Problems scale ($p < .001$).

Table 15 displays the means and standard deviations of the BASC-2 adaptive scale scores for the diagnostic classification groups. MANCOVA were conducted to determine whether the BASC-2 adaptive scales were able to distinguish among diagnostic classification groups when age was controlled.
Table 15

**MANCOVA for BASC-2 Adaptive Scales Using Age (months) as Covariate (ASDs Separated)**

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>AD</th>
<th>PDD-NOS</th>
<th>Other DD</th>
<th>MCHAT Typical</th>
<th>Norm Group</th>
<th>F value</th>
<th>p-value</th>
<th>Partial Eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>40.91</td>
<td>43.69</td>
<td>44.36</td>
<td>50.50</td>
<td>50.47</td>
<td>5.356</td>
<td>&lt;.001</td>
<td>.124</td>
</tr>
<tr>
<td></td>
<td>(9.27)</td>
<td>(11.71)</td>
<td>(12.34)</td>
<td>(7.19)</td>
<td>(10.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Skills</td>
<td>32.52</td>
<td>38.37</td>
<td>34.43</td>
<td>47.74</td>
<td>47.42</td>
<td>22.840</td>
<td>&lt;.001</td>
<td>.375</td>
</tr>
<tr>
<td></td>
<td>(6.19)</td>
<td>(9.70)</td>
<td>(7.08)</td>
<td>(8.28)</td>
<td>(9.18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities of Daily Living</td>
<td>36.74</td>
<td>39.26</td>
<td>43.07</td>
<td>45.09</td>
<td>45.55</td>
<td>3.467</td>
<td>.010</td>
<td>.084</td>
</tr>
<tr>
<td></td>
<td>(9.87)</td>
<td>(12.30)</td>
<td>(12.03)</td>
<td>(12.20)</td>
<td>(9.99)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Communication</td>
<td>34.65</td>
<td>37.71</td>
<td>36.43</td>
<td>45.71</td>
<td>47.05</td>
<td>17.874</td>
<td>&lt;.001</td>
<td>.320</td>
</tr>
<tr>
<td></td>
<td>(5.74)</td>
<td>(9.61)</td>
<td>(4.97)</td>
<td>(7.56)</td>
<td>(8.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Analysis revealed a significant multivariate effect between diagnostic classification groups [Pillai’s Trace = .472, F (32, 588) = 5.09, and p = 0.000]. For the BASC-2 adaptive scales, there was a significant group effect for all four Adaptive scales: Adaptability [F (4, 151) = 5.36, p < .001], Social Skills [F (4, 151) = 22.84, p < .001], Activities of Daily Living [F (4, 151) = 3.47, p = .010], and Functional Communication [F (4, 151) = 17.87, p < .001].

Post-hoc analyses using the Bonferroni technique revealed significant differences between participants in the Autism group and subjects in the M-CHAT Typical (p = .009) and Norm Group (p = .005) on the Adaptability scale, with the Autism group obtaining significantly lower T-scores than either of the typically developing groups. On the Activities of Daily Living scale, subjects in the Autism group scored significantly lower than subjects from the Norm Group (p = .040), but did not demonstrate significant differences from the M-CHAT Typical group on this scale. Participants from the Autism,
PDD-NOS, and Other Developmental Delay groups scored significantly lower on the both the Functional Communication and Social Skills scales than participants from the M-CHAT Typical and Norm Group ($p < 0.001$).

Table 16 displays the means and standard deviations when participants in the Autism group were combined with participants in the PDD-NOS group to reflect all subjects diagnosed with an autism spectrum disorder (ASD). A one-way analysis of variance was used to examine potential differences between participants with an ASD and those with Other Developmental Delays and typically developing subjects on the DSD scale. Significant effects were obtained between groups [$F (3, 154) = 48.068, p < .001$]. Post-hoc tests using the Bonferroni technique revealed there was again a significant difference ($p < .001$) between DSD scores obtained by subjects in the ASD group ($M = 66.43, SD = 9.11$) and those in the typically developing groups (M-CHAT Typical: $M = 54.85, SD = 7.34$; Norm Group: $M = 47.13, SD = 7.43$). There was also a significant difference ($p < .026$) between participants in the ASD group and those diagnosed with Other Developmental Delays ($M = 61.11, SD = 6.86$).

Table 16

*ANOVA for BASC-2 Developmental Social Disorders Scale (ASDs Collapsed)*

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>ASD</th>
<th>Other DD</th>
<th>MCHAT Typical</th>
<th>Norm Group</th>
<th>$F$ value</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSD</td>
<td>66.43 (9.11)</td>
<td>61.11 (6.86)</td>
<td>54.85 (7.34)</td>
<td>47.13 (7.43)</td>
<td>48.068</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The DSD scale was then examined when participants in the ASD group were separated into two diagnostic categories. The means and standard deviations of the Developmental Social Disorders (DSD) scale scores for the Autism, PDD-NOS, Other Developmental Delay, M-CHAT Typical, and Norm Group are presented in Table 17. A one-way analysis of variance was used to examine potential differences between the five groups on the DSD scale. Significant effects were obtained between groups [F (4, 153) = 39.10, p < 0.001].

Post-hoc tests using the Bonferroni technique revealed significant differences between participants in the Autism group and participants classified with other developmental delays (p = 0.001) and those classified as typically developing (p < 0.001), with subjects in the Autism group obtaining significantly higher scores on the DSD scale. Participants in both the PDD-NOS group and those from the Other Developmental Delay group also obtained significantly higher scores on the DSD scale than subjects from the M-CHAT Typical (PDD-NOS = p < 0.001; Other DD = p = .021) and Norm Group (PDD-NOS = p < 0.001; Other DD = p < 0.001). There were no significant differences between participants in the Autism Group (M = 69.74, SD = 7.86) and those in the PDD-NOS group, nor did T-scores from participants in the PDD-NOS (M = 64.26, SD = 9.33) group differ significantly from those classified with Other

Table 17

ANOVa for BASC-2 Developmental Social Disorders Scale (ASDs Separated)

<table>
<thead>
<tr>
<th>BASC-2 Scale</th>
<th>Autism</th>
<th>PDD-NOS</th>
<th>Other DD</th>
<th>MCHAT Typical</th>
<th>Norm Group</th>
<th>F value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSD</td>
<td>69.74</td>
<td>64.26</td>
<td>61.11</td>
<td>54.85</td>
<td>47.13</td>
<td>39.101</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>(7.86)</td>
<td>(9.33)</td>
<td>(6.86)</td>
<td>(7.34)</td>
<td>(7.43)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Developmental Delays (M = 61.11, SD = 6.86). Interestingly, there was a significant difference between DSD scores obtained by subjects in the M-CHAT Typical group (M = 54.85, SD = 7.34) and those from the Norm Group (M = 47.13, SD = 7.43); however, it should be noted that the mean T-score of both groups still fell within the non-elevated range for classification.

**Sensitivity and Specificity of the BASC-2 DSD Scale**

Sensitivity and specificity of the Developmental Social Disorders Scale were derived based on the current sample. Screening effectiveness of T-scores of 60, 65, and 70 were assessed and reported in Table 19. These scores were selected for comparison because according to the BASC-2 manual (Reynolds & Kamphaus, 2004), T-scores between 60-69 represent a score in the “at risk” range, while a T-score of 70 or higher is considered to be in the “clinically significant range.” Therefore, a cut-score of 65 was included in the analysis as a median point between the minimum scores needed for a classification of at-risk or clinically significant. Using a cut score of 60, the Developmental Social Disorders scale accurately screened in 76% of participants identified as having an autism spectrum disorder (i.e., either Autism or PDD-NOS), and accurately screened out 69% of participants without an ASD. A cut score of 65 accurately classified 59% of participants with an ASD while accurately screening out 89% of subjects without an ASD diagnosis. Using a cut score of 70, only 31% of participants with an ASD were screened in, while 97% of participants without an ASD were accurately screened out. As a cut score of 60 resulted in the highest sensitivity, the Positive Predictive Value and Negative Predictive Value of this T-score were also calculated. Utilizing a minimum T-score of 60, the PPV for the DSD scale was .59, while the NPV was .84, resulting in an overall hit rate of .72.
Table 18

*Sensitivity and Specificity of the DSD Scale in Identifying ASD Versus Non-ASD*

<table>
<thead>
<tr>
<th>DSD Cut Score</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>45/58 = 77.59%</td>
<td>69/100 = 69.00%</td>
</tr>
<tr>
<td>65</td>
<td>34/58 = 58.62%</td>
<td>89/100 = 89.00%</td>
</tr>
<tr>
<td>70</td>
<td>18/58 = 31.03%</td>
<td>97/100 = 97.00%</td>
</tr>
</tbody>
</table>

**Correlations between BASC-2 DSD Scale and Other Diagnostic Instruments**

Pearson product moment correlations were conducted to identify associations between the BASC-2 Developmental Social Disorders Scale T-Score and the total scores (Communication + Reciprocal Social Interaction + Imagination/Creativity + Stereotyped Behaviors and Restricted Interests; i.e., \(a + b + c + d\)) of the ADOS and the total score of the CARS, respectively. Correlations were also evaluated when accounting for only the Communication + Reciprocal Social Interactions (i.e., \(a + b\)) domains of the ADOS and the DSD T-score, as well as when Communication + Reciprocal Social Interactions + Stereotyped Behaviors and Restricted Interests (\(a + b + c\)) domain scores were combined. Correlations between the DSD scale score were significant for both the ADOS and CARS scores (see Table 18). The DSD scale score demonstrated a small correlation with the ADOS total score (\(a + b + c + d; r = .254, p < .05\)). Significance of correlations were similar when accounting for only the Communication and Reciprocal Social Interaction scores (\(a + b; r = .250, p < .05\)), and when adding in the score for Stereotyped Behaviors and Restricted Interests (\(a + b + c; r = .223, p < .05\)). The DSD scale score demonstrated a medium to large correlation with the CARS total score (\(r = .439, p < .01\)).
Table 19

*Correlation Matrix of BASC-2 DSD Scale T-Score with ADOS and CARS Total Scores*

<table>
<thead>
<tr>
<th>Measure</th>
<th>ADOS (a + b + c + d)</th>
<th>ADOS (a + b)</th>
<th>ADOS (a + b + d)</th>
<th>CARS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSD T-Score</td>
<td>.254*</td>
<td>.250*</td>
<td>.223*</td>
<td>.439**</td>
</tr>
</tbody>
</table>

* *p < .05
** *p < .01
CHAPTER IV: DISCUSSION

The purpose of the present study was to determine the clinical utility of the BASC-2 Parent Rating Scales, Preschool Form for the differential diagnosis of ASDs in toddler and preschool-age children. As the prevalence of ASDs grows, so does the need for early diagnosis and intervention. As previous researchers (e.g., Paul et al., 2004; Sikora, Hall, Hartley, Gerrard-Morris, & Cagle, 2008; Stone, Ousley, Hepburn, Hogan, & Brown, 1999; Volker et al., 2010) have pointed out, concerns of ASD often first present at primary care and educational settings, which may not be equipped with the time, financial resources, or clinicians with specialized training for comprehensive ASD evaluations. Given these challenges, it may make sense for practitioners such as school psychologists or Head Start consultants to begin an assessment with a broad-based set of rating scales rather than one specific to ASDs (Volker et al., 2010). The BASC-2 is an instrument commonly used in educational and primary care settings when young children evidence behavioral difficulties, and may be of use in identifying toddler and preschool age children at risk for ASDs. The current study examined differences in T-scores on the clinical, adaptive, and Developmental Social Disorders scales between young children diagnosed with ASDs, those diagnosed with other developmental delays, and typically developing children. Validity of the DSD scale in identifying young children with ASDs was also evaluated.

Given the debate regarding the classification of autism spectrum disorders, and the proposed changes to the DSM-V for the classification of ASDs, analyses were
conducted in two ways. First, all participants diagnosed with any ASD were examined as one group. Next, participants in the ASD group were divided into two categories: Autistic Disorder and PDD-NOS, and evaluated separately in comparison to children diagnosed with other developmental delays and the typically developing controls. There are several reasons for conducting two sets of analyses. As stated previously, there is considerable controversy surrounding the current classification of ASDs (Witwer & Lecavalier, 2008; Worley & Matson, 2012). While most researchers agree on the diagnostic classification of Autistic Disorder (Volkmar & Klin, 2005), Witwer and Lecavalier (2008) point out that studies indicate that children diagnosed with PDD-NOS vary widely in their symptom presentation. In their meta-analysis of ASD subtypes, Witwer and Lecavalier (2008) noted that most children diagnosed with PDD-NOS in these studies were observed to exhibit fewer or milder symptoms of the three core ASD subdomains (communication, social interaction, and restricted and stereotyped behaviors) than children diagnosed with Autistic Disorder. However, proposed changes to the DSM-V (American Psychiatric Association, 2011) would eliminate diagnostic subcategories in lieu of a single broad category of ASD. The proposed changes include that at least two symptoms of restricted, repetitive behaviors, interests, or activities (RRB) be present in order to be given a diagnosis of ASD. Children who demonstrate social impairments without the presence of RRBs (including some children who would currently be classified with PDD-NOS) would alternatively receive a diagnosis of Social Communication Disorder (McPartland, Reichow, & Volkmar, 2012). Given the current controversy regarding the validity of ASD subtypes, it was believed to be important to evaluate these participants both separately (i.e., those diagnosed with Autistic Disorder and PDD-NOS) and as a group (all ASDs).
BASC-2 Clinical Scales

The mean scores for participants in the ASD group fell in what authors of the BASC-2 classify as the “at risk” range for both Atypicality and Attention Problems (Reynolds & Kamphaus, 2004). Participants in the ASD group obtained significantly higher T-scores on the Atypicality, Withdrawal, and Attention Problems scales than participants in either of the typically developing groups; however, scores on the these scales did not differ significantly between children with ASDs and those diagnosed with other developmental delays. The same results were obtained when participants diagnosed with autism were evaluated separately from those diagnosed with PDD-NOS, with both groups obtaining significantly higher scores on these scales than participants in the typically developing groups. These results are consistent with findings from previous research (e.g., Knoll, 2008; Mahan & Matson, 2011; Volker et al., 2010). Children in the Other DD group also obtained significantly higher scores on the Atypicality, Withdrawal, and Attention Problems scales than typical controls from the BASC-2 norm sample. Further, differences were not observed between children diagnosed with Autistic Disorder and those with Other DDs on any of the clinical scales when participants from the ASD group were evaluated separately, though subjects diagnosed with PDD-NOS obtained significantly higher T-scores on the Hyperactivity scale than subjects from the norm sample.

Contradictory to the findings by Volker et al. (2010) and Knoll (2008) which found that children and adolescents with high functioning ASDs obtained higher scores on the Depression and Withdrawal scales when compared with typically developing peers, children with ASDs in the current study did not demonstrate these differences. As
seen in Figure 2, participants diagnosed with ASDs in the current study demonstrated notably less impairment across all clinical scales of the BASC-2 than children and adolescents from previous studies. Toddlers and preschoolers in both the autism and PDD-NOS groups obtained mean T-scores in the non-elevated range on both the Depression and Withdrawal scales. One possible explanation for the increase in T-scores on the Depression scale in older children and adolescents is that as children grow older, they may become more aware of differences between themselves and their typically developing peers. Further, teasing and bullying as a result of these differences is likely to increase as children get older, thus leading to increased symptoms of depression (Ghaziuddin, Ghaziuddin, & Greden, 2002; Volker et al., 2010). While the subjects in the previously cited studies were all school-age (i.e., 6-16 years old), participants in the current study were much younger. Although new research (e.g., Luby, 2010) has shown that preschool age children can exhibit depressive symptoms such as loss of interest in play or feelings of guilt, these symptoms are not measured by the Depression scale of the BASC-2. Finally, several items on the Depression scale imply the use of language (e.g., “Says, 'nobody likes me',” “Complains about being teased,” “Is negative”). Given the young age and low T-scores for Functional Communication observed in the current sample, it is likely that many of their parents did not endorse these items.

The young age of the current sample likely contributed to lower T-scores on the Withdrawal scale as well. Given that 7 out of 11 items on the Withdrawal scale involve ratings of interactions with peers (e.g., “Is shy with other children,” “Is chosen last by other children for games”), and most participants in the current sample are not yet
Figure 1. Mean T-scores of clinical, adaptive, and DSD scales by diagnostic classification group.

Figure 2. Mean T-scores of clinical, adaptive, and DSD scales by study population.
school-age, it is possible that many parents in the current sample have not had enough opportunities to observe their child's behavior in the context of same-age peers. Further, items specifically related to friendship (e.g., “Has trouble making new friends,” “Makes friends easily”) may be developmentally inappropriate when examining the behavior of toddlers as children are not likely to begin to develop reciprocal friendships until the preschool period (Patterson, 2008).

**BASC-2 Adaptive Scales**

Congruent with findings from previous research (e.g., Knoll, 2008; Mahan & Matson, 2011; Volker et al., 2010) differences were observed between participants in the ASD group and typically developing children on all of the adaptive scales off the BASC-2. Participants diagnosed with Other Developmental Delays also scored significantly lower than typically developing children on the Social Skills and Functional Communication scales; however, no significant differences were found between participants in the ASD group and those with Other DDs. Consistent with analyses of the BASC-2 clinical scales, no differences were observed between participants diagnosed with autism and those diagnosed with Other DDs when participants from the autism and PDD-NOS group were evaluated separately.

Following the trend observed in the clinical scales, participants diagnosed with ASDs in the current study demonstrated visibly less impairment than children and adolescents from previous studies (see Figure 2). Again, several reasons may account for this discrepancy. First, this study utilized a much younger sample than previous research using the BASC-2. Although the mean age for the current sample was 35 months, the median age was only 30 months, and the mode for the sample was 24 months. As
Zwaigenbaum et al. (2009) points out, toddlers differ from preschool age children in the nature of their social relationships, their cognitive and communicative processes, their learning characteristics, and their daily routines. Even stronger distinctions exist between preschool and school-age children. Therefore there may be less pressure and/or opportunity for toddlers and preschoolers to engage in autonomous behaviors associated with adaptive functioning. Second, parents of younger children are less likely to have an older sibling with which to compare development; thus, ratings of adaptive behavior may portray less impairment because parents are not aware of particular developmental milestones and behaviors observed in typically developing children.

Previous research has noted the utility of including ratings of adaptive behaviors in the assessment of ASDs (Gillham, Carter, Volkmar, & Sparrow, 2000; Perry et al., 2009; Stone, Ousley, Hepburn, Hogan, & Brown, 1999). In a study of young children (ages 22-71 months) comparing children diagnosed with ASDs to those with intellectual disabilities, Perry et al. (2009) found that children with ASDs obtained significantly lower scores on the Communication and Socialization domains of the Vineland Adaptive Behavior Scales. However, no differences were found between children diagnosed with Autistic Disorder and those diagnosed with PDD-NOS on any of the Vineland scales. These results were similar to those obtained by Gillham et al. (2000) with an older sample (ages 4 to 13 years), which found that children with ASDs obtained significantly lower scores on the VABS domains of Socialization and Daily Living Skills than children with other developmental delays. Gillham et al. (2000) also found that children diagnosed with PDD-NOS displayed significantly better communication, socialization, and daily living skills than those diagnosed with Autistic Disorder. These results are
complimentary to those obtained in the current study, which reported that children
diagnosed with PDD-NOS obtained higher scores on the Social Skills adaptive scale than
those diagnosed with Autistic Disorder. However, in the current study, whereas
significant differences were found between children with ASDs and typically developing
children on all of the adaptive scales, none distinguished between children with ASDs
and those with other developmental delays. Still, the ability of the adaptive scales to
discriminate between children with ASDs and typically developing children lends further
support to including examination of adaptive behavior when diagnosing ASDs.

**Developmental Social Disorders Scale**

Examination of the Developmental Social Disorders scale indicated that participants in
Autism, PDD-NOS, and Other DD groups all obtained significantly higher $T$-scores on
the DSD scale than children in either of the typically developing groups (see Figure 3).
Although the mean $T$-score for the Autism, PDD-NOS, and Other DD groups all fell in
what is considered the “at risk” range according to the BASC-2 manual, the mean $T$-score
for participants diagnosed with Autistic Disorder closely approached what would be
considered a clinically significant (i.e., $T \geq 70$) $T$-score (i.e., $M = 69.74$), obtaining
significantly higher scores than participants diagnosed with other developmental delays,
which fell at the low end of the at-risk range (i.e., $M = 61.11$). $T$-scores for subjects
diagnosed with PDD-NOS were observed to be close to the middle of the “at risk” range
(i.e., $M = 64.26$). Mean $T$-scores for each of the typically developing groups were not
elevated. However, there were no significant differences between $T$-scores for
participants in the Autism and PDD-NOS groups, nor were there significant differences
between subjects with PDD-NOS and those diagnosed with Other Developmental Delays.
These results are consistent with previous research (e.g., Trillingsgaard, Sorensen, Nemec, & Jorgensen, 2005; Ventola et al., 2007; Wiggins, 2009; Zwaigenbaum et al., 2009) documenting the difficulty in distinguishing ASDs from other developmental delays in very young children. As previously noted, children with language delays often exhibit impaired social skills, leading to poor social-emotional functioning and behavioral problems, particularly in preschool aged children (Rescorla, Ross, & McClure, 2007; Tervo, 2007). Additionally, some young children with developmental delays may also present with sensory symptoms and repetitive behaviors (Rogers, Hepburn,
Stackhouse, & Wehner, 2003; Ventola et al., 2007), while toddlers with ASDs with more developed language may present with subtler symptoms of ASDs at younger ages.

In two studies comparing children with ASDs to those diagnosed with other developmental delays (i.e., global developmental delay or developmental language delay), both Trillingsgaard et al. (2005) and Ventola et al. (2007) found that deficits in language did not distinguish between the two groups. However, social behaviors such as deficits in joint attention (which are not measured on the DSD scale) were observed more often in children diagnosed with ASDs than in other developmental delays. Wiggins (2009) found that when examining characteristics that distinguish between subgroups of toddlers with ASDs and those with other developmental delays, the severity of autistic symptoms was the best predictor of subgroup classification. The current study also noted the lack of significant differences in scores between participants diagnosed with PDD-NOS and those with Other Dds, thus adding to the hypothesis that toddler and preschool-age children with milder symptoms of ASDs may be more difficult to distinguish from children with other developmental disabilities.

**Sensitivity and Specificity.** Sensitivity and specificity of the Developmental Social Disorders scale was derived based on the current sample. Screening effectiveness of cut scores of 60, 65, and 70 were assessed. The Developmental Social Disorders scale, with a cut score of 60, accurately screened in 78% of participants identified as having an autism spectrum disorder, (i.e., a diagnosis of either Autistic Disorder or PDD-NOS) and accurately screened out 69% of participants without an ASD, including those diagnosed with other developmental delays. Increasing the cut score led to significant increases in specificity, but large reductions in sensitivity. These results are significantly lower than
those obtained by Volker et al. (2010), who found that a cut score of 60 reliably
distinguished children and adolescents with HFASDs from typically developing controls,
with a sensitivity and specificity of .98 and .95, respectively. However, the current study,
while evaluating the scores of a considerably younger group of children, also included
participants with other developmental delays in the calculation of sensitivity and
specificity. As previously noted, ASDs and other developmental delays are particularly
difficult to distinguish in toddler and preschool populations, therefore it is not surprising
that the sensitivity and specificity for this age group is lower. Further, it is expected that
sensitivity and specificity would be much higher in the current sample if participants with
ASDs were only compared to a group of typically developing controls.

The DSD scale accurately classified 21 out of 23 participants diagnosed with
Autistic Disorder; however, it classified 11 out of 34 children diagnosed with PDD-NOS
in the non-elevated range. Studies examining proposed revisions to the criteria for ASDs
in the DSM-V (American Psychiatric Association, 2011), which would require the
presence of restrictive, repetitive, and stereotyped behaviors, have also indicated that
many children diagnosed with PDD-NOS would not meet the criteria for an ASD under
the new symptom definitions (McPartland, Reichow, & Volkmar, 2012; Worley &
Matson, 2012), despite the continued presence of significant impairments. It is possible
that the sensitivity of the DSD scale may increase with implementation of the new
diagnostic criteria of the DSM-V. However, the poor sensitivity of this scale in
identifying children currently diagnosed with PDD-NOS points to the need for greater
awareness of milder symptoms and presentations of ASDs, as well as continued
comprehensive assessment when risk for an ASD is suspected.
**Correlations with Other Measures of ASD.** This study also sought to examine the convergent validity of the BASC-2 Developmental Social Disorders scale with empirically validated measures of autism assessment. Pearson product moment correlations were conducted to identify associations between the BASC-2 Developmental Social Disorders scale T-score and the total scores of the ADOS and the CARS. Results of these analyses found a small, but significant correlation between the T-score on the DSD scale and scores on domains of the ADOS. However, a medium to large correlation was found between the DSD T-score and the total score for the CARS. There are several reasons that may account for this discrepancy. First, the CARS takes into account parental report of behaviors, whereas the ADOS is scored based on behaviors observed by the clinician during administration of the measure. As the DSD scale is based on parent ratings of behaviors, this may also contribute to the stronger correlation between T-scores on the DSD scale and the total score of the CARS. Second, items on the CARS measure a broader range of behaviors than what is evaluated on the ADOS, some of which may better correspond to items on the DSD scale. For example, the item “has a short attention span” from the DSD scale is similar to “activity level” as rated on the CARS, “adjusts well to changes in routine,” another DSD item, may correlate with “adaptation to change” as measured on the CARS, while “throws tantrums” (DSD) can be correlated with ratings of “emotional response” on the CARS. Although these behaviors may all be observed during administration of the ADOS, and in fact the ADOS allows for the coding of “overactivity” and “tantrums, aggression, negative or disruptive behavior,” these ratings are not included in the final diagnostic algorithm for classification of ASDs, and thus were not included in the current analysis.
Implications for Practitioners

Findings from the current study have several implications for practitioners working with toddlers and preschoolers. First, while more research is needed to replicate results and make definitive recommendations, preliminary evidence suggests that examining BASC-2 profiles of young children referred for behavioral or developmental concerns may aid in identifying those at risk for having an ASD. However, clinicians should be aware of the low specificity of the DSD scale observed in the current sample of children who presented with milder forms of ASDs. Therefore, elevations in other scales such as Atypicality, Attention Problems, Social Skills, and Functional Communication should also be considered risk factors. Practitioners are further cautioned against using such rating scales in isolation for making diagnostic decisions, but rather as a starting point to guide further assessment.

Strengths, Limitations, and Directions for Future Research

The current study builds on existing literature by focusing on participants in the toddler and preschool population. This is especially important considering the emphasis on early diagnosis and intervention. Also unique to this study was the inclusion of a clinical sample of children diagnosed with other developmental delays (Global Developmental Delay and Developmental Language Delay) in addition to the typically developing control group. Additionally, stringent criteria were followed for including children in the clinical samples. Rather than relying on parental self-report on special education eligibility, all children in the clinical samples (including those diagnosed with other developmental delays) were diagnosed via comprehensive evaluations administered by trained clinicians, including gold-standard measures of autism assessment.
Despite its strengths, several limitations exist with the current study and should be addressed in future research. First, the present study utilized a clinical sample of children who were referred for an evaluation because they failed an autism screening instrument. Thus, results obtained from participants in the ASD, other DD, and M-CHAT Typical groups may not be representative of children with corresponding diagnoses in the general population. Second, subject characteristics such as ethnicity, socio-economic status, nonverbal IQ, and language ability were not included in the current analyses.

With regard to the clinical samples used in the current study, participants diagnosed with Global Developmental Delay and Developmental Language Delays were collapsed into a single diagnostic group, as there were not enough subjects with each of these diagnostic classifications to be analyzed separately. It is possible that more differentiation may exist between children with ASDs and those with other developmental delays when these diagnostic categories are disaggregated.

A final limitation of the current study is that differences in the specific diagnostic criteria used to diagnose children with Autistic Disorder or PDD-NOS was not evaluated. For example, the DSM-IV-TR criteria for a diagnosis of Autistic Disorder specifies that a child must exhibit a total of six or more items from each of the behavioral categories (i.e., qualitative impairment in social interaction, qualitative impairment in communication, and restricted, repetitive, and stereotyped patterns of behavior), with at least two symptoms related to impairments in social interaction, and one each from impairments in communication and restricted, repetitive, and stereotyped behaviors. Alternatively, a diagnosis of PDD-NOS requires impairment in social interactions as well as impairments in communication or the presence of stereotyped behavior, interests, and activities.
Future studies should look at the specific symptoms endorsed when diagnosing participants that may account for differences in behavioral profiles.

In addition to examining the specific symptoms and diagnostic endorsed when comparing subtypes of ASDs, there are a number of directions for future research. First, replications of the current study using larger clinical and control groups are needed before the BASC-2 DSD scale can be endorsed as a valid screening instrument for ASDs. Second, research comparing BASC-2 profiles of children with ASDs to those diagnosed with other disorders (such as Disruptive Behavior Disorders) would be valuable in discriminating between challenging behaviors observed in children with ASDs versus patterns seen in children with other disorders. Finally, research should be conducted with educational and primary care agencies to determine if the use and examination of behavior rating scales such as the BASC-2 result in a difference in diagnostic outcomes and service delivery for young children at risk for ASDS.
References


http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6103a1.htm?s_cid=ss6103a1_w


development of an infant with autism from birth to two years of age. *Journal of Applied Developmental Psychology*, 21(3), 299-313. doi:10.1016/S0193-3973(99)00042-8


Autism Spectrum Disorders With the Social Communication Questionnaire.

*Journal of Developmental and Behavioral Pediatrics*, 27(Suppl2), S95-S103.
doi:10.1097/00004703-200604002-00007


Shevell, M., Ashwal, S., Donley, D., Flint, J., Gingold, M., Hirtz, D., …Sheth, R. D.


South, M., Williams, B. J., McMahon, W. M., Owley, T., Filipek, P. A., Shernoff, E. E.,


doi:10.1023/A:1005683209438


Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., ...Yirmiya, N.
Table A1

**DSM-IV-TR Criteria for Autistic Disorder**

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

1. **Qualitative impairment in social interaction, as manifested by at least two of the following:**
   a. marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction
   b. failure to develop peer relationships appropriate to developmental level
   c. a lack of spontaneous seeking to share enjoyment, interests, or achievements with others (e.g., by a lack of showing, bringing, or pointing out objects of interest)
   d. lack of social or emotional reciprocity

2. **Qualitative impairments in communication as manifested by at least one of the following:**
   a. delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)
   b. in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
   c. stereotyped and repetitive use of language or idiosyncratic language
   d. lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level

3. **Restricted, repetitive, and stereotyped patterns of behavior, interest, and activities, as manifested by at least one of the following:**
   a. encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
   b. apparently inflexible adherence to specific, nonfunctional routines or rituals
   c. stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movements)
   d. persistent preoccupation with parts of objects

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

C. The disturbance is not better accounted for by Rett's Disorder or Childhood Disintegrative Disorder.
Table A2.

*BASC-2 PRS-P Items Contributing to DSD Scale*

<table>
<thead>
<tr>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Compliments others</td>
</tr>
<tr>
<td>6. Has a short attention span</td>
</tr>
<tr>
<td>9. Has trouble making new friends</td>
</tr>
<tr>
<td>30. Provides full name when asked</td>
</tr>
<tr>
<td>43. Communicates clearly</td>
</tr>
<tr>
<td>54. Makes friends easily</td>
</tr>
<tr>
<td>64. Bangs head</td>
</tr>
<tr>
<td>73. Acts strangely</td>
</tr>
<tr>
<td>75. Encourages others to do their best.</td>
</tr>
<tr>
<td>78. Is chosen last by other children for games.</td>
</tr>
<tr>
<td>97. Adjusts well to changes in routine.</td>
</tr>
<tr>
<td>98. Shows feelings that do not fit the situation.</td>
</tr>
<tr>
<td>117. Throws tantrums.</td>
</tr>
</tbody>
</table>