2017

# The Effect of Demographic and Clinical Factors on the Performance of the Abbreviated Scoring Algorithm for the Baby and Infant Screen for Children with Autism Traits

Paige Elizabeth Cervantes
*Louisiana State University and Agricultural and Mechanical College*, pcerva2@lsu.edu

# THE EFFECT OF DEMOGRAPHIC AND CLINICAL FACTORS ON THE PERFORMANCE OF THE ABBREVIATED SCORING ALGORITHM FOR THE BABY AND INFANT SCREEN FOR CHILDREN WITH AUTISM TRAITS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctorate of Philosophy

In

The Department of Psychology

By
Paige E. Cervantes
B.A., Binghamton University, 2012
M.A., Louisiana State University, 2015
August 2017

## TABLE OF CONTENTS

## ABSTRACT

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction and social communication as well as the presence of restricted, repetitive, and stereotyped patterns of behavior, interests, or activities (RRBIs; American Psychiatric Association [APA], 2013). Individuals with ASD experience lifelong and pervasive impairments across many domains of functioning. Early identification of ASD is imperative as earlier intervention is associated with greater gains in numerous areas. In regards to early identification efforts, routine screening is important. However, screening often fails to be conducted within early childhood care settings. To encourage greater rates of screening by offering a time efficient measure with sound psychometric properties, previous researchers developed and validated the abbreviated scoring algorithm for the *Baby and Infant Screen for Children with aUtIsm Traits, Part 1 (BISCUIT-Part 1)*. The current study further examined the utility of the abbreviated scoring algorithm as a screener specifically in regards to its performance across varying ages, genders, and levels of developmental functioning. Results indicated that the abbreviated scoring algorithm performed well without any changes to its original form across age groups and genders and in participants with typical developmental functioning. An increase in the cutoff score was necessary to achieve adequate sensitivity and specificity in identifying ASD risk for participants with low developmental functioning. Implications of these results are discussed.

# CHAPTER 1. INTRODUCTION

Once thought to be a rare condition, autism spectrum disorder (ASD) is now known to affect approximately 1% of the general population (Elsabbagh et al., 2012; Fombonne, Quirke, & Hagen, 2009; Matson & Kozlowski, 2011; Newschaffer et al., 2007). Individuals with ASD face lifelong difficulties with pervasive impairments in social communication, social interaction, and restricted, repetitive behaviors and interests (RRBIs; Landa & Garrett-Mayer, 2006; Lord & Luyster, 2006; Newschaffer et al., 2007). However, due in part to greater professional and public autism awareness, interventions have been established to effectively mediate some of the difficulties experienced by individuals with ASD (Eldevik et al., 2009; Virués-Ortega, 2010). Specifically, early intensive behavioral intervention (EIBI) using the principles of applied behavior analysis (ABA) has been shown to produce, on average, robust and comprehensive treatment effects in children with autism (Darrou et al., 2010; Kamio, Haraguchi, Miyake, & Hiraiwa, 2015; Virués-Ortega, 2010). EIBI, however, has been found to produce the best outcomes in the youngest participants; therefore, it can be assumed that earlier enrollment in an intervention program would be beneficial (Granpeesheh, Dixon, Tarbox, Kaplan, & Wilke, 2009).

Because of the effectiveness of early enrollment in EIBI, early ASD identification is of the utmost importance and has been at the forefront of research and clinical efforts for some time (Wiggins, Baio, & Rice, 2006; Williams, 2006). Although researchers consistently demonstrate that a reliable ASD diagnosis can be made as early as 2 years of age (Baird et al., 2001; Daniels & Mandell, 2013; Guthrie, Swineford, Nottke, & Wetherby, 2013; Johnson, Myers, & and the Council on Children With Disabilities, 2007; Lord & Luyster, 2006), a substantial portion of children with autism are not diagnosed until 3 to 4 years old (Daniels & Mandell, 2013; Matson,

Wilkins, & Gonzalez, 2008; Wetherby et al., 2004). Further, there continues to be a significant time lag between first parental concern and age of diagnosis as well as an average one year long delay between the first developmental evaluation and the initial diagnosis of ASD. This situation leads to increased parental stress and missed opportunity to enroll in appropriate intervention services early on (Wiggins et al., 2006).

In order to improve early detection of ASD, the field has advocated for the performance of routine autism screening. Several measures to detect autism risk in infants and toddlers have therefore been developed (Wiggins et al., 2006). These screening tools likely help expedite the process of receiving an autism diagnosis. Screeners are typically used in primary care or early intervention settings to identify at-risk children; immediately following a positive screen, these children are often referred to a specialty clinic for a more comprehensive evaluation (Campbell, Scheil, & Hammond, 2016). However, recommended screening procedures often fail to be carried out (Crais et al., 2014; Gillis, 2009; Johnson et al., 2007; Matheis & Matson, 2015). The low rate of screenings completed in primary care and early intervention settings highlights the importance of the development and modification of tools so that they are both psychometrically strong and convenient and easy to administer in practice. As autism is a very heterogeneous disorder, it is also necessary that these tools demonstrate utility over a range of ASD phenotypes (Wiggins et al., 2006).

To provide an efficient method for screening within Louisiana's statewide early intervention program, Cervantes, Matson, and Peters (in press) recently developed an abbreviated scoring algorithm for the *Baby and Infant Screen for Children with aUtIsm Traits, Part 1 (BISCUIT-Part 1)*. The abbreviated scoring algorithm is made up of six items, has adequate psychometric properties, and shows promise in its ability to function as a brief

screening tool. However, more research is necessary. Therefore, the aim of the current study was to further examine the clinical utility of the algorithm to serve as a screener. The effect of various clinical and demographic factors on the performance of the algorithm in detecting autism risk in infants and toddlers was evaluated. Specifically, age, gender, and level of developmental functioning were explored.

## CHAPTER 2. AUTISM SPECTRUM DISORDER (ASD)

### History of ASD

Though autism is currently understood as a neurodevelopmental disorder characterized by deficits in social skills and communication, and RRBIs, the term was first used in the early 1900's by Eugen Bleuler, a Swiss psychiatrist, to describe specific symptoms of a different condition. Bleuler (1913) used "autistic thinking" as an antonym of rational thinking to describe the turning away from reality evident in those affected by schizophrenia. In fact, many children with ASD symptoms were diagnosed with childhood schizophrenia at the time. It was not until 1943 that Leo Kanner introduced and conceptualized the disorder we now know as autism. In his paper titled "Autistic Disturbances of Affective Contact," Kanner (1943) described the similar symptoms observed in 11 children that could not be categorized into an existing disorder. Though these children (eight males and three females) differed in family and developmental history, degree of impairment, and symptom manifestation, Kanner noted that each child presented with communication deficits, a desire for the maintenance of sameness, and substantial socialization impairments. Kanner felt that an "extreme autistic aloneness," or an inability to interact with other people, was most central to this condition. Kanner termed this condition Infantile Autism in his 1944 follow-up study.

In regards to communication deficits, three of the 11 children examined in Kanner's study never developed speech. The remaining eight children were verbal; however, their language was often nonfunctional. Children with language often engaged in echolalia (i.e., repeating words or phrases that were previously heard) and pronoun reversal (i.e., referring to oneself using the incorrect pronoun such as "he," "she," or "you"). Further, the children with more grammatically correct language in Kanner's sample demonstrated unusual and selective

responding in conversations with others; these children frequently responded only to topics related to their restricted interests (Kanner, 1944). Socially, Kanner noted that the children he studied isolated themselves and often resisted interactions with others. Kanner theorized that these pervasive social impairments were present in these children since birth. Last, all of the children in Kanner's sample demonstrated a desire to maintain sameness in their environment. The children were described as repetitious and followed very strict and ritualized routines. Kanner reported that the children became very disturbed at any interruptions in their routines, and that they were limited in their ability to engage in a variety of spontaneous behaviors (Kanner, 1943).

Coincidentally, Hans Asperger of Austria published an account of a very similar symptom set in children at the same time as Kanner (Asperger, 1991). In his paper "Autistic Psychopathy in Childhood," Asperger detailed the behavioral presentations of four children who differed extensively within several domains but each presented with a common "autistic" personality. Among characteristics shared within these children, Asperger noted that social integration difficulties were paramount. Symptoms also exhibited by the children in Asperger's sample included communication issues, stereotypic movements, rigidity, pica, and conduct problems. Like Kanner, Asperger also emphasized symptom persistence over time. Though his accounts may have added strength to Kanner's distinction of autism as a standalone disorder, Asperger's paper did not receive wide recognition until 1991 when it was translated into English by Uta Frith (Asperger, 1991).

Despite much early work in researching and distinguishing autism from similar childhood disorders (e.g., schizophrenia, intellectual disability [ID]; Kanner, 1943, 1944; Lockyer & Rutter, 1970; Rutter & Bartak, 1971; Rutter, 1968; Rutter & Schopler, 1987), the *Diagnostic and*

*Statistical Manual of Mental Disorders (DSM)* did not include autism for nearly four decades after Kanner's initial description. It was not until the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition* (*DSM-III*; APA, 1980; Rutter & Schopler, 1987) that autism was recognized in its text. From 1980 until 2013, autism was conceptualized as an "umbrella" disorder consisting of several unique subgroups captured under one diagnostic category in all *DSM* editions (e.g., *DSM-III*, *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised* [*DSM-III-R*], *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* [*DSM-IV*], and *Diagnostic and Statistical Manual of Mental Disorders*, *Fourth Edition, Text Revised* [*DSM-IV-TR*]; APA, 1980, 1987, 1994, 2000). The diagnostic category was labeled pervasive developmental disorders (PDDs). The term PDDs was very carefully chosen in order to differentiate autism from mental illness (e.g., schizophrenia). As such, emphasis was placed on the developmental features of the disorder and on the range of behavioral domains affected (e.g., social, communication, behavior; Rutter & Schopler, 1987). PDD subcategories evolved with *DSM* changes over time. Once consisting of Infantile Autism, Atypical Autism, and Childhood Onset PDD in the *DSM-III*, five separate subcategories more recently formed PDDs in the *DSM-IV-TR*. These included Autistic Disorder, Asperger's Disorder, Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS), Rett's Disorder, and Childhood Disintegrative Disorder.

For purposes of better understanding changes made in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*), *DSM-IV-TR* definitions for Autistic Disorder, Asperger's Disorder, and PDD-NOS follow. Autistic Disorder was diagnosed when an individual presented with six or more of the diagnostic items (i.e., ≥ two items from the social domain, ≥ one item from the communication domain, and ≥ one from the RRBI domain). Social

domain items consisted of (1) impairment in nonverbal behaviors (e.g., eye contact, gestures, facial expression); (2) difficulty developing relationships with peers; (c) lack of sharing enjoyments, interests, or achievements with others; and, (4) failure to engage in social or emotional reciprocity. Communication domain items were (1) lack of or delayed spoken language; (2) difficulty initiating and maintaining conversation; (3) stereotyped and repetitive use of language or idiosyncratic language; and, (4) limited socially imitative or pretend play skills. The RRBI domain included (1) preoccupation with one or more stereotyped or restricted interests; (2) strict adherence to specific, nonfunctional routines; (3) stereotyped and repetitive motor movements; and, (4) preoccupation with parts of objects. In addition, individuals must have exhibited symptoms prior to 3 years of age for an Autistic Disorder diagnosis (APA, 2000).

Asperger's Disorder was diagnosed in the *DSM-IV-TR* if an individual exhibited symptoms within the social and RRBI domains described above but had no impairment in language, cognitive, or adaptive functioning. PDD-NOS was given when an individual evinced social deficits and demonstrated communication impairments or RRBIs. PDD-NOS is similar to Autistic Disorder in presentation but the full Autistic Disorder criteria is not met due to either late age of onset or a subthreshold presentation of symptoms (APA, 2000).

## Diagnostic Criteria

Following concerns regarding the adequacy of the *DSM-IV-TR* criteria, the APA released the *DSM-5* in 2013; this new edition presents several significant changes to the autism definition (APA, 2013; Gibbs, Aldridge, Chandler, Witzlsperger, & Smith, 2012; Matson, Kozlowski, Hattier, Horovitz, & Sipes, 2012; Worley & Matson, 2012). First, the authors of the *DSM-5* collapsed the Autistic Disorder, Asperger's Disorder, and PDD-NOS subcategories into one single ASD category. This change was made due to possible ambiguity in the boundaries

between autism subcategories and in the diagnostic criteria of Asperger's Disorder and PDD-NOS (Lord & Bishop, 2015). The diagnoses of Rett's Disorder and Childhood Disintegrative Disorder are removed from ASD. To help differentiate subgroups within this broad ASD diagnosis, several specifiers relating to severity, presentation, and comorbidity are provided in the *DSM-5*. Severity specifiers are assigned based upon the amount of supports needed for an individual to function. Level 1 ASD describes an individual who is at the highest level of functioning but "requires support." Without support, symptoms related to social communication and social interaction will cause noticeable impairments presenting as lack of back and forth conversation, unsuccessful attempts to make friends, or atypical responses to social cues. RRBIs in individuals with Level 1 ASD interfere with functioning in at least one context. A severity specifier of Level 2 indicates that a person "requires substantial support." Individuals with Level 2 ASD have marked impairments in both verbal and nonverbal social communication, and social skill deficits are apparent even with supports in place. In regards to RRBIs, individuals with Level 2 ASD may show inflexibility in their behavior, difficulty coping with change, and RRBIs that are apparent and interfere with functioning in at least one context. Level 3 ASD describes individuals with the most severe symptomology and who require "very substantial support." Individuals with Level 3 ASD have severe impairments in communication and social skills (e.g., limited speech, little to no initiation of and response to social interactions), severe behavior inflexibility, severe difficulty coping with change, and RRBIs that markedly interfere in almost all contexts of functioning (APA, 2013).

When making a *DSM-5* ASD diagnosis, a clinician is also encouraged to use neurobiological specifiers to indicate the co-occurrence of a medical or genetic condition (e.g., epilepsy, Fragile X syndrome), or an environmental factor (e.g., very low birth weight).

Specifiers for accompanying intellectual or language impairment are also available. Unique to the *DSM-5*, modifiers (i.e., other behavioral diagnoses that reflect comorbidity) can accompany a diagnosis of ASD. For example, ADHD and ASD could not be dually diagnosed in the *DSM-IV-TR*; however, ADHD can now serve as a modifier for ASD in the *DSM-5* (APA, 2013; Lord & Bishop, 2015).

Specific criteria for ASD changed as well. Due to research evidence indicating *DSM-IV-TR* communication domain items were not distinct from social domain items, the social and communication domains were merged into a single diagnostic domain (i.e., Social Communication and Social Interaction; Lord & Bishop, 2015). Within this Social Communication and Social Interaction domain, individuals must exhibit symptoms from all three items to qualify for diagnosis. Items include (1) deficits in social-emotional reciprocity, (2) nonverbal communication impairments, and (3) difficulty in developing, maintaining, and understanding relationships. To be eligible for a diagnosis in the *DSM-5,* an individual must meet at least two of the four symptoms in the RRBI domain. This change has increased from one of four RRBI items in the *DSM-IV-TR*. RRBI domain items in the *DSM-5* consist of (1) stereotyped, repetitive motor movements, use of objects, or speech, (2) insistence on sameness, strict adherence to routines, or ritualized patterns of verbal or nonverbal behavior, (3) restricted, fixated interests that are abnormal in intensity or focus, and (4) atypical reactivity to sensory input or unusual interest in sensory aspects of the environment. Age of onset requirements also changed from the *DSM-IV-TR* to the *DSM-5*. While the *DSM-IV-TR* required symptoms to be present before 3 years old (APA, 2000), the *DSM-5* more vaguely states that ASD symptoms must be exhibited in early childhood but "may not fully manifest until social demands exceed limited capacities" (APA, 2013).

Great controversy in the field of autism followed the release of the *DSM-5* regarding such large changes in a long accepted definition of ASD (Frazier et al., 2012; Gibbs et al., 2012; Matson, Hattier, & Williams, 2012; Matson, Kozlowski, et al., 2012; McPartland, Reichow, & Volkmar, 2012; Wing, Gould, & Gillberg, 2011). With the tightening of diagnostic criteria from the *DSM-IV-TR* to the *DSM-5* described above, approximately 40% of individuals with a *DSM-IV-TR* ASD diagnosis have been found to no longer meet criteria for autism in the *DSM-5* (Frazier et al., 2012; Matson, Kozlowski, et al., 2012; Mattila et al., 2011; McPartland et al., 2012; Smith, Reichow, & Volkmar, 2015; Worley & Matson, 2012). At heightened risk for no longer meeting diagnostic criteria are individuals with ASD and IQs over 70, individuals with PDD-NOS, and individuals with Asperger's Disorder (Mattila et al., 2011; Smith, Reichow, et al., 2015; Worley & Matson, 2012). Thus, the change in diagnostic criteria may lead to a more affected ASD population.

Multiple studies have been conducted to compare groups of individuals meeting only *DSM-IV-TR* criteria and individuals with a *DSM-5* ASD diagnosis, and consistent evidence was made for a *DSM-5* ASD population with more developmental delays, more frequent and severe challenging behavior, and more severe autism symptomology (Beighley et al., 2013; Beighley, Matson, Rieske, Konst, & Tureck, 2014; Turygin, Matson, Beighley, & Adams, 2013; Worley & Matson, 2012). The introduction of the *DSM-5* may also have implications for early identification and diagnosis of ASD. Researchers have posed that because more and more severe ASD symptoms are necessary to qualify for a *DSM-5* diagnosis and ASD symptoms become more prominent with age, infants and toddlers may not reach the diagnostic threshold when professionals use a strict application of the new criteria. Children may need to be older before

ASD can be diagnosed using the *DSM-5* holding serious consequences for the availability of appropriate early intervention services (Zander & Bölte, 2015).

## Prevalence of ASD

Autism was once thought to be a rare condition with original prevalence estimates at four to five individuals per 10,000 (Lotter, 1966; Matson & Kozlowski, 2011). However, following more intensive empirical investigation beginning in the1990s, prevalence estimates for autism were substantially higher. This rise from original estimates sparked concern and speculation of an epidemic (Inglese & Elder, 2009). Subsequently, the Center for Disease Control and Prevention (CDC) instated the Autism and Developmental Disorders Monitoring Network (ADDM) to investigate and monitor the prevalence of autism and related developmental disabilities (DDs) around the United States. The first ADDM study began in 2002, and ASD prevalence estimates were examined in 8-year-old children over 14 locations across the United States. This first study resulted in a prevalence rate of one in 150 children affected by ASD, a male to female gender ratio between 3.4:1 and 6.5:1, and higher rates of ASD in non-Hispanic white children (CDC, 2007).

Since the first ADDM investigation, the prevalence of ASD has shown an increasing trend. The most recent ADDM study was published in 2014 and resulted in a prevalence estimate of one in 68 children affected by ASD, a 29% increase from the previous ADDM study published in 2008 and a 123% increase from the first ADDM study (CDC, 2014). The male to female ratio remained constant at approximately 4.5:1; however, there were significantly more children with ASD and average or above average cognitive functioning than previously found (i.e., 32% in 2002 versus 46% in 2010; CDC, 2014). In terms of racial and ethnic discrepancies, non-Hispanic white children were 30% more likely than black children and nearly 50% more

likely than Hispanic children to be identified with ASD (CDC, 2014). However, differences in ASD across racial/ethnic groups have not been replicated across studies and may be more attributable to access to quality care and/or relevant cultural factors.

The CDC offers a cost-efficient, population-based protocol for estimated ASD prevalence. Their statistics are widely used to describe the rise in autism. However, researchers and clinicians have reported several methodological concerns that warrant attention; these include flaws in standardization of diagnosis, lack of direct observation and assessment of children, and wide variation in results across sites. Rather than reflect true occurrence or non-occurrence of autism, results from the record review conducted as part of their protocol may reflect the rate at which clinicians and/or educators assess and document signs of ASD (Mandell & Lecavalier, 2014). Therefore, it would be erroneous to treat the estimates derived from the ADDM studies as true prevalence rates. However, in culminating the findings across prevalence studies from various research teams using different methodologies, ASD is assumed to affect around 0.6 - 1% of the population (Elsabbagh et al., 2012; Fombonne et al., 2009; Newschaffer et al., 2007).

Therefore, estimated prevalence rates for ASD across research groups are much higher than once suspected. Definitive explanations for this increase are less clear, and it is more likely that a number of variables are contributing to the rise in ASD cases (Matson & Kozlowski, 2011). Perhaps the most prominent factor in these changes is the ever-changing definition of ASD. Since Kanner's first conceptualization of autism, the diagnostic criteria for ASD have broadened substantially to include a wider range of individuals (Fombonne, 2003; Matson & Kozlowski, 2011; B. Taylor, 2006). Greater acceptance, awareness, and service availability for ASD contributes to the increasing prevalence as well, as it has helped encourage earlier

identification and diagnosis (Fombonne, 2003; Leonard et al., 2010; Matson & Kozlowski, 2011). This increased awareness in the general public may also trigger increased misdiagnosis of ASD. Evaluating an individual for autism is a complex task due to heterogeneity in symptom presentation as well as a great deal of shared symptoms with other DDs; thus, with increased attention on ASD, some autism diagnoses at present may be inaccurately assigned (Leonard et al., 2010; Matson & Kozlowski, 2011). Further, diagnostic substitution (i.e., replacing a former diagnosis with a different diagnosis) is likely contributing to the rise in ASD cases. A common example of diagnostic substitution involves the historic conceptualization of ID. Years ago, a person with ID and symptoms consistent with autism was typically given a primary diagnosis of ID. Due to increased knowledge and understanding of the ASD and ID relationship, a primary diagnosis of ASD and a secondary diagnosis of ID would be a more accurate formulation for the abovementioned case today (Leonard et al., 2010). In a similar fashion, cultural factors related to the understanding of ASD as a disorder contribute to ASD prevalence estimates. Additionally, the varying research methods used to calculate rates may influence the ASD prevalence rates (Fombonne, 2003; Leonard et al., 2010; Matson & Kozlowski, 2011).

Of note, various environmental factors speculated to relate to the rise in autism cases have received great public attention. Among these, autism has been linked to factors such as intolerance to certain foods and exposure to several types of infections, medications, vaccinations, and toxins. However, these variables have yet to be supported in the research literature for contributing to the ASD increase (Inglese & Elder, 2009). The proposed link between autism and vaccinations (particularly the measles, mumps, rubella [MMR] vaccine and vaccines containing thimerosal) has been most notable in the realm of environmental factors due to the substantial media coverage and great controversy that followed. Andrew Wakefield

sparked this vaccination controversy when he published his 1998 study linking the MMR vaccination to the onset of autism symptoms in eight children. Though this study was extremely methodologically flawed and eventually retracted from the journal in which it was published, speculation about the link grew as did public concern (B. Taylor, 2006). With further research, evidence has become increasing clear that there is no association between autism and vaccines (L. E. Taylor, Swerdfeger, & Eslick, 2014); however, consequences of this speculation remain (Johnson et al., 2007). Not only do a significant amount of parents of children with autism believe their child's ASD was caused by immunizations, many new parents are choosing not to vaccinate their children due to fear of autism; as a result, the frequency of measles outbreaks has grown internationally (Phadke, Bednarczyk, Salmon, & Omer, 2016; B. Taylor, 2006; L. E. Taylor et al., 2014).

On the contrary, pre- and perinatal factors and the continuing improvement of prenatal and neonatal care are likely contributing to the rise in ASD cases. For example, the survival rate for children with low birth weight and prematurity rises with enhanced medical care; and, researchers have found that infants with low birth weight and/or who are born premature are at a two-fold increased risk for autism (Matson & Kozlowski, 2011; Schendel & Bhasin, 2008). Therefore, there may be a multitude of factors associated with the increased prevalence of ASD.

## Etiology of ASD

Though autism was first conceptualized more than five decades ago, its cause is still unknown. Etiological studies are complicated due to large genetic complexity and variation in how ASD presents between and even within individuals over time (Johnson et al., 2007). Early etiological theories emphasized the importance of parent qualities in causing autism.

Emotional coldness, obsessive traits, and lack of affection on the part of the child's parents, and particularly of the mothers, were often implicated in autism (Rutter, 1968). In fact, Bruno Bettelheim (1967) proposed a "refrigerator mother" theory suggesting parent hostility and wish that their child did not exist led to the emergence of autism symptoms in children. Other theories at the time included autism symptoms as a result of the child's attempts to mask an alternate existing psychopathology (i.e., schizophrenia), of aberrant reticular system activity, and of a lack of ability to understand sound (Bender, 1959; Rutter, 1968). There is no empirical support for these early claims; instead, it can be assumed that autism etiology is heterogeneous and involves a complex interaction between multiple genetic factors and, to a lesser extent, the environment (Cheslack-Postava & Jordan-Young, 2012; Johnson et al., 2007).

In regards to the genetic and neurobiological underpinnings of ASD, there is overwhelming evidence against the idea that a single brain dysfunction or genetic sequence exists to explain all cases of autism (Cheslack-Postava & Jordan-Young, 2012; Johnson et al., 2007; Waterhouse & Gillberg, 2014). Across genetic studies, variations among genes or loci on nearly every chromosome have been implicated in the etiology of ASD (Cheslack-Postava & Jordan-Young, 2012). With advanced technology for examining genes, researchers have also found certain copy number variations (i.e., insertions or deletions of segments of DNA) that may be associated with the emergence of ASD symptoms; however, research findings have been conflicting (Cheslack-Postava & Jordan-Young, 2012). While exact genetic influence remains unknown in most cases of autism, a minority of ASD cases (< 10%) can be linked to genetic syndromes. Individuals with a known syndrome or medical condition that is causitive or associated with autism (e.g., Fragile X syndrome, Tuberous Sclerosis, Phenylketonuria, Fetal Alcohol Syndrome, Angelman Syndrome) are sometimes subtyped as having "secondary ASD"

(Cheslack-Postava & Jordan-Young, 2012; Johnson et al., 2007). These individuals often evince comorbid intellectual impairments and are more likely to have dysmorphic features (Johnson et al., 2007). "Idiopathic ASD" refers to individuals who have autism but do not have a co-occuring and associated genetic or medical condition. This is not to say that idiopathic ASD is not biologically based but rather multiple genes and gene interactions are likely involved in its etiology. To demonstrate the influence of genetics, the recurrence risk of idiopathic ASD is approximately 5% when there is an older sibling with ASD; this rate is significantly higher than the risk within the general population (Johnson et al., 2007). In addition to genetics, neurological studies have found numerous atypicalities within the brain structure and function of the ASD population. Studies suggest increased brain volume, gray- and white-matter differences in several brain regions (e.g., frontal, limbic, basal ganglia, cerebellum), possibly impaired brain connectivity, and abnormal head growth within the first years of life are present in many individuals with autism. People with ASD have also been found to show abnormalities in areas of brain activity during the processing of certain information (e.g., facial recognition, executive functioning; Johnson et al., 2007; Waterhouse & Gillberg, 2014).

There are also many environmental factors theorized to play into etiology and influence autism presentation. Advanced parental age at conception (both maternal and paternal) has been linked to increased risk for autism (Cheslack-Postava & Jordan-Young, 2012; Johnson et al., 2007). Various prenatal variables have also been associated with autism including maternal illness during pregnancy and exposure to certain medications in utero. Perinatally, birth weight, gestational age, and trauma at the time of birth have all been proposed risk factors for autism (Cheslack-Postava & Jordan-Young, 2012; Johnson et al., 2007). Though much speculation has

occurred for postnatal factors related to autism, less empirical evidence has been found to support these (e.g., the previously mentioned vaccination link; Johnson et al., 2007).

Looking forward, research on etiology should and will continue. Environmental risk factors should be further explored. In addition, our understanding of neurobiological causal factors will grow with continued advancements in genetic and neuroimaging technology. In regards to future research, Waterhouse and Gillberg (2014) suggest changing our approach. The researchers believe we need to stop looking for a unitary explanation for the cause of autism. Instead, we should subtype ASD beyond idiopathic and secondary groups based upon presentation and clinical characteristics to explore the possibility of multiple etiological models (Waterhouse & Gillberg, 2014). Continued research is important, as achieving a better understanding of autism etiology may lead to the development of more accurate screening and diagnostic approaches.

# CHAPTER 3. ASD IN INFANTS AND TODDLERS

## Studying Symptoms of ASD in Young Children

Developing an understanding of autism onset and presentation in infant and toddlerhood is imperative for improving ASD screening procedures in young childhood. For several decades, various retrospective and prospective strategies have been used to study early ASD emergence. Retrospective strategies have included caregiver report and the study of home videos; prospective strategies follow several at-risk populations through early development to evaluate symptom emergence. This research helps in the development and validation of early screening measures and procedures.

### Retrospective Studies

One of the most common methods for acquiring information regarding early symptom emergence in children with ASD is parent report. This method is used both in research and in practice. Gathering information from parent report has several advantages including efficiency and familiarity with child behavior across a wide range of settings and time (Ozonoff, Heung, Byrd, Hansen, & Hertz-Picciotto, 2008; Zwaigenbaum et al., 2007). Several limitations to parent report exist. Caregiver report is subject to memory errors, restrictions in the ability to notice more subtle impairments and developmental atypicalities, and confirmatory biases. Parents are more likely to confirm early symptoms consistent with their child's current diagnosis in their report (Cervantes, Matson, & Goldin, 2016; Ozonoff et al., 2008; Rogers, 2009).

Another retrospective strategy employed is analyzing home videos (Rogers, 2009; Saint-Georges et al., 2010; Zwaigenbaum et al., 2007). Studying home videos of children who are diagnosed with ASD allows for a more objective evaluation of symptom emergence by trained viewers; however, some weaknesses remain (Zwaigenbaum et al., 2007). Limitations include

concerns regarding video representativeness of the child's behavior across all contexts and the standardization of videos for study when families do not videotape their children for the same duration or during the same events (Cervantes et al., 2016; Ozonoff et al., 2008; Rogers, 2009; Saint-Georges et al., 2010; Zwaigenbaum et al., 2007).

**Prospective Studies**

Prospective studies use experimental methods to examine the emergence of ASD in a standardized fashion (Zwaigenbaum, 2010). Prospective studies make use of high-risk populations (e.g., children with siblings with ASD, children who fail population screeners, children who have medical/genetic diagnoses that frequently co-occur with ASD) and involve multiple assessments over a long duration of time. These high-risk infants and toddlers are compared to low-risk children on standardized developmental measures and naturalistic evaluations throughout their young lives ending most often at age 3 years old (Landa & Garrett-Mayer, 2006). A large advantage of prospective designs is that they are not subject to the same reporting biases as in retrospective designs. Also, data collection methods across participants and time can be made more uniform (Landa & Garrett-Mayer, 2006; Rogers, 2009).

Prospective studies can be limited however in their age of enrollment and age at endpoint. The later the age of enrollment, the higher risk for sampling biases to be in effect, as parents may be aware of their child's aberrant behaviors consistent with ASD (Rogers, 2009). Age of endpoint of the study also holds implications. The earlier in development a study ends, the more likely children with milder presentations of ASD are misclassified as typically developing (Landa & Garrett-Mayer, 2006; Rogers, 2009). Milder forms of ASD are usually not diagnosed until after 3 years old when social demands increase and symptoms become more pronounced. Limitations also exist specific to the type of high-risk population used. Within

sibling studies, the severity of the older sibling's ASD may influence participation; generalization may also be limited due to possible differences between simplex and multiplex families (Landa & Garrett-Mayer, 2006; Zwaigenbaum, 2010). Because studies using children who fail developmental screeners cannot begin until participants are old enough to be reliably assessed with a screener, typical age of enrollment is rarely under 1 year of age. This allows for potential sampling biases to occur because parents may already have concerns (Zwaigenbaum, 2010). Lastly, studying children with conditions that are often comorbid with ASD restricts the generalization of results; the conditions that are used are relatively rare, and these children often have unique autism phenotypes (Cervantes et al., 2016; Landa & Garrett-Mayer, 2006).

## Symptoms of ASD in Infants and Toddlers

In regards to autism onset, researchers have shown most parents become increasingly aware of ASD symptoms or lack of developmental progression within the first 2 years of their child's life (Werner, Dawson, Munson, & Osterling, 2005). Marked differences in both social communication and RRBIs between young children who are typically developing, developmentally delayed, and who later receive a diagnosis of ASD are evident in the assessment of many children around 2 years old (Wetherby, Watt, Morgan, & Shumway, 2007; Zwaigenbaum, 2010).

Over half of parents of children with ASD report atypicalities with language development as their first concern (De Giacomo & Fombonne, 1998). Symptoms noted as early as 12 months of age include limited or lack of language, abnormal use of language, difficulties with receptive language, and limited use of communicative gestures (Chawarska, Klin, Paul, & Volkmar, 2007; Rogers, 2009; Saint-Georges et al., 2010; Wetherby et al., 2004; Zwaigenbaum, 2010). Though onset of social impairments in many cases occurs several months after birth, parents appear to be

less sensitive to socio-emotional atypicalities in their young children (De Giacomo & Fombonne, 1998). However, infants and toddlers who are later diagnosed with ASD often show limited social interest and responsiveness, difficulties interacting with others, aberrations in emotions and facial expressions, deficits in eye gaze and eye contact, less social smiling, impairments in play skills, decreased sharing of interests, deficiencies in imitation, and a lack of or limited joint attention skills (Saint-Georges et al., 2010; Wetherby et al., 2004; Zwaigenbaum, 2010; Zwaigenbaum et al., 2007).

RRBIs are less indicative of ASD in infanthood; however, distinguishing symptoms within this domain often emerge between the first and second year of life in children who are diagnosed with ASD (Saint-Georges et al., 2010; Wetherby et al., 2004). RRBIs in early development often include non-functional object use, unusual exploration of objects, and hyper- or hypo-reactivity to sensory input (Rogers, 2009; Zwaigenbaum, 2010). In regards to assessing repetitive motor movements and differentiating those that are typical in development from those that are characteristic of ASD, children with autism often show atypical persistence, quality, and frequency of behavior compared to children without autism (Zwaigenbaum et al., 2007).

## Patterns of ASD Symptom Emergence

As previously stated, most young children who are later diagnosed with ASD experience a gradual onset of symptoms starting very early on in development and continuing to grow in prominence with age. However, there is a subset of individuals with ASD who have a late onset of ASD symptoms. Approximately 30% of children who are diagnosed with ASD experience regression (Johnson et al., 2007; Ozonoff et al., 2008). Regression involves a loss of previously acquired skills that can be related to a number of developmental domains and that typically occurs before 3 years old. Children who experience regression most often show both social and

language skill loss (Kalb, Law, Landa, & Law, 2010; Ozonoff et al., 2008).

It was previously thought that autistic regression solely involved a degradation of functioning following typical development; however, this is no longer the case. Though instances have been noted where children are not distinct from typically developing peers at their first birthday but show no symptom differences from early onset ASD cases by their second, there are several other onset patterns that may be more common (Ozonoff et al., 2008; Saint-Georges et al., 2010). One such pattern is referred to as a "developmental plateau." Children who experience this onset show mostly typical developmental progression until age 2 years old when the progression stops. Researchers believe this pattern reflects an inability of children with ASD to build upon basic social skills to achieve greater complexity in their social repetoire (Kalb et al., 2010; Ozonoff et al., 2008; Saint-Georges et al., 2010; Tager-Flusberg, 2010). Another onset pattern that deviates from the traditional conceptualization of regression involves the demonstration of subtle developmental discrepancies prior to a loss of skills. This is theorized to be the most frequent onset pattern and is characterized as a "mixed onset" (Ozonoff et al., 2008).

Conflicting findings have been presented on the implications of varying onset patterns for ASD presentation and prognosis. Some researchers have found worsened communication, social, and behavioral prognoses for children who have late onset ASD while others found no differences in presentation or prognosis between early and late onset cases (Kalb et al., 2010; Werner et al., 2005). However, the idea that autism onset can occur later in a subset of children is very widely accepted. Therefore, conducting standardized screenings at several timepoints across infant and toddlerhood is integral for catching these children with later onsets and improving early ASD identification (Kalb et al., 2010; Ozonoff et al., 2008; Rogers, 2009).

## CHAPTER 4. EARLY IDENTIFICATION AND THE ROLE OF SCREENING

Improving early identification procedures has been at the forefront of autism research for some time. This is because early ASD diagnosis allows for earlier initiation of appropriate treatment and medical care, needed family supports (e.g., genetic counseling, caregiver psychoeducation, stress management), and planning for future educational options (Baird et al., 2001; Crais et al., 2014; Filipek et al., 2000). With the continued rise in professional and public awareness of autism, greater service availability and treatment effectiveness, and improved diagnostic tools, advancements in early identification and diagnosis are likely. However, there are some concerning trends found in the early identification research that require remediation.

First, an average 1 to 2 year time lag has been found between age of first parental concern and age at diagnosis (Crais et al., 2014). Further, De Giacomo and Fombonne (1998) found that while parents had concerns beginning around 19 months of age, a mean of 5 months passed before they reported their concerns and sought the advice of professionals. This gap may be due to limited parental knowledge regarding not only autism but also how typical development occurs in young childhood. Parents or caregivers may also be quick to dismiss the seriousness of their early concerns and attribute atypicalities to problems their child will grow out of in time (De Giacomo & Fombonne, 1998). This perception would delay seeking consultation.

When concerns are acted upon, primary care physicians (PCPs) are usually the first to field parental concerns (De Giacomo & Fombonne, 1998; Johnson et al., 2007; Wetherby et al., 2004). However, researchers have found that PCPs may not be proficient or knowledgeable in identifying and understanding early autism symptoms (De Giacomo & Fombonne, 1998; Goin-Kochel, Mackintosh, & Myers, 2006; Shah, 2001). Lack of understanding of autism may limit

PCPs ability to appropriately address parental concerns regarding their child's development (Cervantes et al., 2016; Goin-Kochel et al., 2006; Shah, 2001).

On average, parents report visiting around four to five clinicians on the way to an ASD diagnosis, and approximately 40% of parents of children with autism report being unsatisfied with the process (Goin-Kochel et al., 2006). Delayed diagnosis is detrimental for several reasons; first, it is unlikely that young children and their families can obtain effective early intervention services without a diagnosis or an at-risk label (Crais et al., 2014). Because EIBI initiated before the age of 4 years old has been linked to superior language, social, and cognitive improvements, this is particularly concerning (De Giacomo & Fombonne, 1998). Additionally, delayed diagnosis has been shown to take a toll on family wellbeing and increase stress. Further, parents may respond inappropriately to child behavior when atypicalities are not well understood making later treatment more difficult (Baird et al., 2001). Because earlier diagnosis betters the prognosis for both children with ASD and their families, screening procedures should be considered an integral component within the early identification framework. Formal ASD screening in early childhood care settings may help remediate many of the barriers discussed above.

## CHAPTER 5. ASD SCREENING PROCEDURES

As the first contact for families with concerns regarding their child's development, PCPs need to have both the ability to identify early signs of autism and a strategy for systematic evaluation of developmental progression (Johnson et al., 2007). To aid in systematic assessment for autism, the American Academy of Pediatrics (AAP) recommends both universal surveillance and screening of children by PCPs. Surveillance entails continued monitoring of child development through eliciting and attending to parent concerns, maintaining a developmental history, observing and interacting with the child, and documenting findings. Surveillance procedures should be carried out at every visit through childhood (Johnson et al., 2007). However, PCP evaluation of development based upon clinical judgment alone has been found to be less accurate than formal screening tools (Crais et al., 2014; Johnson et al., 2007). Because of this, the AAP recommends a general developmental screening tool be administered at the 9-, 18-, and 30-month visits to identify developmental delay and a universal ASD screening be conducted at the 18- and 24-month visits to identify risk for autism. The AAP advocates for universal ASD screening at two time points to help identify children who may have experienced a late onset pattern. ASD screening should also be conducted whenever concerns consistent with ASD symptoms are raised (Barton, Dumont-Mathieu, & Fein, 2012; Johnson et al., 2007).

Following a positive screen on an ASD screener, PCPs should be able to specify next steps for the family to take in the evaluation process. These clinicians must be knowledgeable about local resources in order to make an appropriate referral for a more specialized and intensive diagnostic assessment (Barton et al., 2012; Gray & Tonge, 2005; Johnson et al., 2007). However, specialized diagnostic clinics often have long waitlists and can be difficult to access.

Therefore, it is recommended that PCPs also refer at-risk children to early intervention programs to ensure treatment be initiated as soon as possible (Barton et al., 2012).

## Arguments Against Universal Screening

Most of the arguments against universal screening center around the notion that we lack effective screening instruments and evidence-based interventions for autism (Al-Qabandi, Gorter, & Rosenbaum, 2011; Mandell & Mandy, 2015). Screening tools are often over-inclusive in order to maximize the rate of true positive screens; this results in a heightened rate of false positive screens. Contributing to the high rate of false positives are the screening of children with DDs other than ASD and the screening of children with early developmental concerns that later resolve as a variation of typical development (Barton et al., 2012; Johnson et al., 2007). Those who oppose universal screening frequently report concerns regarding the unnecessary stress for parents and waste of resources for follow-up evaluation resulting from a false positive (Oosterling et al., 2010; Williams, 2006). Opponents of universal screening also note the difficulty of getting service needs met following a true positive screen. Resources must be available to refer clients to for comprehensive evaluation, early intervention services, parent education and family support, and care coordination services (Dosreis, Weiner, Johnson, & Newschaffer, 2006). Other concerns include the ethics behind diagnosing individuals in the absence of parental concern, the stigma behind a true positive, and the harmful effects of false negative screens (e.g., provides a false sense of reassurance to caregivers, can delay diagnostic and treatment services; Oosterling et al., 2010; Williams, 2006).

**US Preventative Services Task Force (USPSTF) Recommendation Statement and Clinician Rebuttals**

In early 2016, the US Preventative Services Task Force (USPSTF) published a recommendation statement related to autism screening that went against the recommendations made by the AAP. Due to an alleged lack of "evidence about the balance of benefits and harms," the USPSTF recommended against the practice of universal ASD screening (Mandell & Mandy, 2015; Robins et al., 2016; Siu et al., 2016). The harms of screening noted by the USPSTF included the potential for misdiagnosis as well as the time and anxiety associated with further evaluation following false positive screens (Siu et al., 2016). Though the USPSTF reported that available screening tools can detect ASD reliably in young children and that there is quality evidence for the effectiveness of EIBI for autism (Mandell & Mandy, 2015; Siu et al., 2016), the committee believed there to be "inadequate direct evidence" of the benefits of screening in children for whom no concerns consistent with autism had been reported by caregivers or PCPs. The committee also had concerns regarding the lack of studies focusing on outcomes of young children identified with ASD through screening (Siu et al., 2016).

More specifically, the USPSTF was concerned about the generalizability of early intervention studies as most recruited participants were from specialist ASD services rather than participants with ASD who had been detected through screening; the USPSTF said these participants identified through screening may be younger in age and less affected (Mandell & Mandy, 2015; Robins et al., 2016). Several researchers have offered rebuttals for this concern. First, screening rarely leads directly to treatment. Instead, children are referred to a specialty clinic for a more comprehensive evaluation prior to the initiation of intervention services (Mandell & Mandy, 2015). Next, there is no evidence that children with ASD who are identified

through screening will be less likely to respond to early intervention. In fact, it is more likely that these children will be younger at the start of treatment, and there is a large amount of evidence demonstrating that younger individuals show a greater response to treatment. Further, children with more mild impairments have been found to show the greatest responsiveness to treatment as well; though, there is no indication that the children detected through screening are more mildly affected (Mandell & Mandy, 2015; Robins et al., 2016).

The USPSTF also reported the need for greater research in order to recommend the practice of universal ASD screening. The committee details a needed costs and benefits assessment and required longitudinal studies of children who screened negative to test the specificity of available screening tools long-term (Mandell & Mandy, 2015). Though this research would be beneficial, the necessary studies reported by the USPSTF are both time-intensive and financially expensive. If we wait for such studies to be completed before engaging in routine screening, autism will likely continue to go unidentified in many young children, particularly those of low socioeconomic status and/or of a minority race or ethnicity (Mandell & Mandy, 2015; Robins et al., 2016). Therefore, this recommendation statement presented by the USPSTF can have harmful effects. For one, policy makers may decrease efforts to facilitate universal screening and insurance companies may stop reimbursing such services. If, as a result, early childhood professionals stop screening for ASD, children with ASD and perhaps other DDs that go undetected may miss opportunities for early initiation of appropriate and effective intervention services and family supports (Mandell & Mandy, 2015; Robins et al., 2016).

## Benefits of Universal Screening

Though the debate remains active in regards to universal screening, there are several advantages to this approach. In fact, many clinicians believe early diagnosis outweighs the

potential negative effects of screening (Barton et al., 2012). As previously mentioned, surveillance alone does not reliably detect autism symptoms (Crais et al., 2014; Johnson et al., 2007; Robins et al., 2016). Further, parents may not share their concerns if not directly asked; this is especially true of parents who are not proficient in the English language (Hyman & Johnson, 2012; Johnson et al., 2007). Relatedly, lack of parent concern regarding child behavior and development does not directly indicate that the child is typically developing. Therefore, reliance on formal screening measures is integral in early identification for some cases (Johnson et al., 2007).

Further, routine screening offers a more systematic method for monitoring and attending to children's developmental abnormalities. Caregivers of children with ASD have reported dissatisfaction with the way their concerns were handled by their child's PCP and noted that they would have preferred faster referral for specialty evaluation and intervention over the reassurance and monitoring they received (Hyman & Johnson, 2012). Formal screening procedures provide a framework where concerns are more likely to be attended to and referrals are more likely to be made when necessary (Hyman & Johnson, 2012). Researchers have shown that screening often leads to earlier diagnosis of ASD as well as earlier contact with intervention services (Oosterling et al., 2010).

One of the biggest concerns raised by opponents of universal screening are the harmful effects caused by false positives (e.g., stress, wasted time and resources in further evaluation). However, a false positive screen is often linked to symptoms of a related but separate DD (e.g., global developmental delay [GDD], language disorder; Barton et al., 2012; Crais et al., 2014). Therefore, referring false positive children for more extensive assessment is valuable due to the ability to achieve differential diagnosis. In fact, many parents prefer the opportunity to detect a

developmental delay and intervene early even if the positive screen results in greater worry (Barton et al., 2012). Comprehensive and specialized evaluation of children who screen positive on ASD screeners but do not have ASD may help identify another DD, determine if the child is eligible for early intervention, and ensure more thorough monitoring of development (Crais et al., 2014).

## Barriers to Screening

Despite the benefits, routine ASD screening often fails to be conducted within early childhood care settings (Crais et al., 2014; Gillis, 2009; Johnson et al., 2007; Matheis & Matson, 2015). In fact, while over 40% of PCPs reported having at least 10 patients with ASD, only 8% endorsed conducting routine ASD screens in their practice (Johnson et al., 2007). These rates of routine ASD screening in primary care settings have been consistently estimated under 30% in the research literature (Crais et al., 2014). Further, screening has been found to be significantly less likely to occur with clients of a minority race or ethnicity in primary care settings (Arunyanart et al., 2012).

Several barriers to screening have been identified that help explain the low rate of ASD screens being conducted. First, PCPs report a lack of familiarity with and comfort in using screening tools (Crais et al., 2014; Dosreis et al., 2006; Gillis, 2009). Some concerns regarding the accuracy of these screening measures have also been raised (Hyman & Johnson, 2012). Another common problem PCPs note is a lack of training related to child development and autism specifically (Crais et al., 2014; Gillis, 2009). Time constraints are frequently reported by clinicians in primary care settings as well (Crais et al., 2014; Dosreis et al., 2006; Gillis, 2009; Hyman & Johnson, 2012). Further concerns include insufficient reimbursement for performing these services. Insurance companies rarely reimburse clinicians for the whole screening process;

the time needed for interpreting the results of the screen and providing feedback and direction to parents following screening is often not covered (Dosreis et al., 2006; Hyman & Johnson, 2012). Lastly, many professionals report having limited knowledge in how to follow-up a screen. Inadequate referral resources or not knowing where to refer to help clients obtain a more comprehensive diagnostic evaluation are frequently cited issues (Crais et al., 2014; Hyman & Johnson, 2012). Therefore, efforts should be made in the development of both new tools and procedures to help overcome these barriers to screening.

## ASD Screening Tools for Infants and Toddlers

The screening tools used in assessment of ASD risk are generally categorized as Level 1 or Level 2. Level 1 screening tools are used for the widespread screening of a population for ASD. These tools are most often used within primary care settings to distinguish between children who are at risk of ASD and those without risk. Level 1 tools are typically quick to administer and score and are easy to use (Barton et al., 2012; Johnson et al., 2007). Level 2 screening tools are used to distinguish between children at risk for ASD from children at risk for other DDs (e.g., GDD, language disorder) in developmental clinics, early intervention programs, or as part of a larger diagnostic assessment battery. These measures are typically more time intensive and require greater training to administer, score, and interpret (Johnson et al., 2007).

The ideal ASD screening tool would be both psychometrically strong and convenient for use in busy healthcare settings. In regards to psychometrics, four dimensions are typically taken into account when judging the quality of a screening tool (Barton et al., 2012). First, sensitivity estimates should be high. Sensitivity is the ability of the screener to correctly identify children who are at-risk for ASD. Specificity estimates should be considered as well and describe the ability of the screener to correct identify those without risk for ASD. Last, positive predictive

value (PPV) and negative predictive vale (NPV) warrant attention in selecting a screening tool. PPV measures the proportion of children who screen positive on the screener that have ASD; and, NPV measures the proportion of children who screen negative on the screener and do not have ASD. A screener has psychometric strength when it reliably produces specificity, sensitivity, PPV, and NPV estimates at or above 0.80 (Campbell et al., 2016). Of note, NPV is often high and PPV is often low even with adequate sensitivity and specificity estimates when screening for low prevalence conditions like ASD (Barton et al., 2012; Johnson et al., 2007). Screening tools are also designed to maximize true positives and capture as many at-risk children as possible, sometimes at the expense of PPV (e.g., more false positives; Campbell et al., 2016; Johnson et al., 2007). In addition to having sound psychometric properties, screening tools need to be designed for use in settings where screening is not always a top priority service and is not fully reimbursed. Therefore, the screener must be time efficient in administration, scoring, and interpretation, accessible and affordable, and easy to use so that it can be integrated into existing practices (Barton et al., 2012). In the following section, the properties of available autism screening tools are described.

**CHecklist for Autism in Toddlers (CHAT)**

The *CHecklist for Autism in Toddlers* (*CHAT*; Baron-Cohen, Allen, & Gillberg, 1992) was the first screening tool designed to detect autism risk in 18-month olds within the general population. The *CHAT* assesses autism symptomology in two ways, through a nine-item parent report measure and a five-item PCP observation. Key areas of assessment are related to social interest, different aspects of joint attention (e.g., protoimperative and protodeclarative pointing), motor development, and play (i.e., rough and tumble play, social play, pretend play, functional play; Baron-Cohen et al., 1992). In the first study of the *CHAT*'s psychometrics, performance of

the screener was tested on 50 low-risk infants from the general population during routine check-ups and 41 high-risk infants who had an older sibling with ASD. The *CHAT* correctly identified the four high-risk children who went on to later be diagnosed with ASD. Further, the measure did not identify any of the typically developing siblings of children with ASD and did not identify any of the low-risk infants without ASD as having risk for autism (Baron-Cohen, 1992).

A second study on the psychometric strength of the *CHAT* attempted to replicate the promising 1992 results with an expanded the sample size ($n$ = 16,000 18-month-olds; Baron-Cohen et al., 1996). Based upon scores on an initial *CHAT* screen, participants were assigned to varying risk groups including an autism risk group (i.e., failed all five "protodeclarative pointing," "gaze-monitoring," and "pretend play" items), a developmental delay risk group (i.e., failed one or two key items: "pretend play," or "protodeclarative pointing" and "pretend play"), and a normal group; the administration of comprehensive diagnostic evaluations followed this group assignment for a minority of the participants. According to results of the evaluations, ten of the 12 children in the autism risk group received an autism diagnosis and the two children who did not have autism received a diagnosis of developmental delay. When reassessed over three years later, diagnoses were found stable. Of the 22 children in the developmental delay risk group, 15 were diagnosed with developmental delay, seven had no diagnosis, and no children in this group were found to have autism (Baron-Cohen et al., 1996; Campbell et al., 2016).

To establish sensitivity, specificity, and PPV estimates for the *CHAT*, Baird and colleagues (2000) studied the tool's performance for identifying autism risk in 16,235 18-month-old children over the course of 6 years. Using the high-risk group criteria defined above, the *CHAT* had a sensitivity of 0.20, a specificity of 0.998, and a PPV of 0.289. Using the

developmental delay risk group criteria discussed prior (referred to in this study as "medium-risk" criteria), the sensitivity estimate increased to 0.38, the specificity was 0.98, but the PPV decreased to 0.05 (Baird et al., 2000). The authors also studied a potential two-phase administration where children who scored at risk on the first administration of the *CHAT* were screened again 1 month later. Using the high-risk criteria, PPV increased to 0.75; however, sensitivity was reduced to 0.18. In using the medium-risk criteria with a two-phase administration, PPV was 0.294 and sensitivity was 0.20. Specificity was over 0.99 using both high-risk and medium-risk methods (Baird et al., 2000). Therefore, the *CHAT* was not able to identify a majority of the children at risk for autism. Poor sensitivity estimates and high rates of false negative screens as well as the time needed to administer both parent report and observation items warrant concern regarding effective and convenient clinical use (Campbell et al., 2016).

**The Modified CHecklist for Autism in Toddlers (M-CHAT)**

The *Modified Checklist for Autism in Toddlers (M-CHAT*; Robins, Fein, Barton, & Green, 2001) is a Level 1 screening tool developed as an extension to the *CHAT* (Baron-Cohen et al., 1992). This measure includes 23 parent report items, nine of which were retained from the *CHAT,* that require a yes/no response. Six critical items of the *M-CHAT* were identified and include "protodeclarative pointing," "response to name," "interest in peers," "bringing things to show parents," "following a point," and "imitation." Children who fail three of the 23 items overall or two of the six critical items are considered at-risk for ASD (Campbell et al., 2016; DiGuiseppi et al., 2010; Robins et al., 2001). The authors of the *M-CHAT* identified reliability estimates using a two-pronged approach; parents of children who scored at-risk on the screener were contacted for a telephone interview to clarify items failed. Then, based upon the results of the telephone interview, a group of participants were invited for a full diagnostic evaluation.

Participants included both children screened during routine checkups with their PCP at 18- or 24-months (i.e., low-risk; *n* = 1,122) and children screened through their early intervention program (i.e., high-risk; *n* = 171). Of the 1,293 total participants, 1,161 children passed the *M-CHAT*, 74 participants required a telephone interview but did not need a comprehensive evaluation, 19 children were evaluated and diagnosed with language or global delay but not autism, and 39 were evaluated and diagnosed with ASD. None of the children that were identified as needing further evaluation by the *M-CHAT* were found to have typical development (Robins et al., 2001). Because all of the participants had not yet received a follow-up evaluation at the time of publication, Robins and colleagues (2001) offered psychometric estimates rather than absolute values. Sensitivity estimates ranged from 0.87-0.97, specificity estimates ranges from 0.85-0.99, PPV ranged from 0.36-0.80, and NPV was estimated at 0.99. Further, internal reliability was adequate at 0.85 for the entire measure and 0.83 for the critical items only (Robins et al., 2001).

The psychometric properties of the *M-CHAT* were studied again in 2006 with a sample of 84 2- to 3-year-olds referred to an autism clinic by a PCP (Eaves, Wingert, & Helena, 2006). The sensitivity, specificity, and PPV values were estimated for both the critical item algorithm and the measure as a whole. Sensitivity for the six critical items was estimated at 0.77 and specificity at 0.43; PPV was 0.65. For the complete measure, sensitivity was 0.92, specificity was 0.27, and PPV was 0.68 (Eaves et al., 2006). Of note, the sample used in this study consisted of a majority of children who were later diagnosed with ASD (*n* = 54; 64.29%) and children who were older than the sample in the study done by Robins and colleagues (2001). Eaves and colleagues (2006) also used different cutoff criteria and a different formula to determine PPV compared to the 2001 study (Kleinman et al., 2008; Robins et al., 2001).

Because of these methodological differences and the need for further study, Kleinman and colleagues (2008) sought to calculate PPV and reliability estimates on 3,309 low-risk children (i.e., screened during routine checkup with PCP) and 484 high-risk children (i.e., screened through early intervention program or referred from developmental pediatrician or psychologist). Internal reliability was found to be 0.85 and 0.84 for the 23-item *M-CHAT* and the six critical items, respectively. PPV was estimated at 0.11 for the low-risk sample and 0.60 for the high-risk sample. The high rate of false positives was improved with the application of the follow-up telephone interview; the PPV for the low-risk group rose to 0.65, and the PPV for the high-risk group rose to 0.76 (Kleinman et al., 2008). Kleinman and colleagues (2008) also conducted a second screening and evaluation procedure with clients initially screened between 16-30 months old at around 4 years old. Results from this procedure indicated that 76 of the 201 participants that failed the initial *M-CHAT* were later diagnosed with ASD (PPV=0.38); further, 73 of the 124 children who failed the combined initial screening and telephone interview were diagnosed with ASD at 4 years old (PPV=0.59). In regards to false negatives, seven children who were diagnosed with ASD at the follow-up evaluation did not fail the initial *M-CHAT* screen. These children were not significantly different than children who failed the screen and received an ASD diagnosis on age, gender, autism symptoms, or adaptive skills (Kleinman et al., 2008).

In one of the largest and most recent studies on the performance of the *M-CHAT*, Chebowski, Robins, Barton, and Fein (2013) studied 18,989 low-risk children aged 16-30 months old. Of the total sample, 1,737 children failed the initial screen; 79% of those children were followed up with a telephone interview, and 272 continued to screen positive. One hundred seventy-one children were then given a comprehensive evaluation. Though only 54% of these

children were diagnosed with ASD following evaluation, 98% met criteria for a *DSM* diagnosis or were labeled developmentally delayed (Chebowski et al., 2013). Similar to the Kleinman and colleagues (2008) study, the PPV of the *M-CHAT* without subsequent application of the follow-up interview was low (0.06). Therefore, the follow-up interview appears crucial for the adequate performance of the *M-CHAT* (Campbell et al., 2016).

The *M-CHAT* has been translated into several different languages, is used internationally, and has been subject to several revisions (Robins et al., 2014; Robins & Dumont-Mathieu, 2006). For example, researchers have demonstrated the improved effectiveness of a recent revision to the *M-CHAT*, the *Modified CHecklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F*; Robins et al., 2014). The *M-CHAT-R/F* is a two-stage screener. The first stage includes 20 yes/no parent report items retained from the original *M-CHAT*. If a child fails at the first level, assessors (e.g., PCPs, early intervention providers) ask parents structured follow-up questions to obtain additional information such as examples of at-risk behaviors at the second stage (Robins et al., 2014). Internal reliability was found at 0.79 for the *M-CHAT-R/F*. Similar to the results of previous studies, children who met or surpassed cutoff scores for the initial and follow-up assessments were found to have a near 50% risk of being diagnosed with ASD and a 95% risk of having developmental delays (Robins et al., 2014). In sum, results appear mixed in regards to the psychometric properties of the *M-CHAT,* particularly when used without application of a follow-up interview. However, the measure is currently the most commonly used and studied ASD screening tool available (Campbell et al., 2016).

**Pervasive Developmental Disorders Screening Test-II (PDDST-II)**

The *Pervasive Developmental Disorders Screening Test, Second Edition (PDDST-II;* Siegel, 2004*)* is designed for use with children aged 12-48 months old and offers a parent report

autism screening system. As such, the *PDDST-II* is made up of three forms to be used within different settings: Stage 1 - the Primary Care Screener (PCS); Stage 2 - the Developmental Clinic Screener (DCS); and, Stage 3 - the Autism Clinic Severity Screener (ACSS). The PCS consists of 22 items for parents to respond either "yes, usually true" or "no, not usually true." Using a sample of approximately 650 children at-risk for ASD and 250 children with risk for other DDs (i.e., children who were born preterm and expected to show atypical development but not autism), Seigel (2004) calculated a sensitivity of 0.92 and a specificity of 0.91 based upon agreement between item and group classification (Zwaigenbaum & Stone, 2006). Therefore, because research into the PCS used only high-risk participants and was based on classification into groups defined by clinical suspicion rather than clinical diagnoses, further study is required to validate its utility within a variety of settings (Campbell et al., 2016; Dumont-Mathieu & Fein, 2005; Zwaigenbaum & Stone, 2006).

The DCS consists of 14 items intended to diferentiate children at risk for autism from children at risk for other DDs. On a sample of 490 children with ASD and 194 diagnosed with other DDs, sensitivity and specificity estimates of the DCS were found to be 0.73 and 0.49, respectively (Siegel, 2004). However, Zwaigenbaum and Stone (2006) noted some limitations in how these estimates were calculated including a high base rate of autism in the sample. The ACSS is made up of 12 items designed to predict severity of ASD symptoms. In comparing ACSS scores of children with ASD and children with PDD-NOS or Asperger's Disorder, sensitivity was estimated at 0.58 and specificity was estimated at 0.60 (Campbell et al., 2016; Siegel, 2004). Overall, the psychometrics found for the PCS are stronger than for the DCS and ACSS; however, limitations in the methodological procedures used by the authors denote the

need for greater study on the psychometric properties and utility of all *PDDST-II* forms (Campbell et al., 2016; Dumont-Mathieu & Fein, 2005; Zwaigenbaum & Stone, 2006).

**Screening Tool for Autism in Two-year-olds (STAT)**

The *Screening Tool for Autism in Two-year-olds* (*STAT*; Stone, Coonrod, & Ousley, 2000) was designed to screen for ASD in children aged 24-35 months old. The *STAT* is unique because it is an interactive assessment measure. The 12 items of the *STAT* reflect four distinct domains of social-communication behavior including play skills, motor imitation, requesting, and directing attention. Each item is scored as pass or fail according to criteria specified within the manual. Domain scores reflect the proportion of failed items to total items, and the total score is achieved by summing the four domain scores (Stone et al., 2000; Stone, Coonrod, Turner, & Pozdol, 2004). The *STAT* takes approximately 20 minutes to administer and requires clinician training.

Stone and colleagues (2000) sampled 73 children with ASD or another DD to develop and validate an optimal *STAT* scoring algorithm. Psychometric characteristics were examined; sensitivity was estimated at 0.83, specificity was 0.86, PPV was 0.77, and NPV was 0.90 (Stone et al., 2000). A follow-up study was conducted in 2004 (Stone et al., 2004) with a sample of 52 children who had either a clinical diagnosis of ASD or a developmental delay or language impairment. Using a cutoff score of 2, the *STAT* resulted in a sensitivity estimate of 0.92, specificity of 0.85, PPV of 0.86, and NPV of 0.92. Five children without ASD had false positive screen results, and one child with ASD had false negative screen results (Stone et al., 2004). Both inter-observer agreement and test-retest reliability estimates were found to be high. Further, evidence of concurrent validity was found through high agreement with the *Autism Diagnostic Observation Schedule-General (ADOS-G;* Stone et al., 2004).

Recent research has shown that the *STAT* is not only a promising screening tool for 2-year-olds but also can be used to reliably detect autism risk in children under 2 (Stone, McMahon, & Henderson, 2008). A sample of 71 children aged 12-23 months with either an older sibling with ASD or who were referred for evaluation due to developmental concerns related to autism was examined to determine the performance of the *STAT* on younger children (Stone et al., 2008). Participants were screened with the *STAT* at the initial evaluation and then provided a comprehensive diagnostic assessment at a follow-up evaluation. Results indicated that utilizing the cutoff score previously developed on older children was inadequate (i.e., sensitivity of 1.0 but a specificity of 0.40). Therefore, the authors recommended an increase in the cutoff for younger samples (total score $\geq 2.75$). With the new optimal cutoff score, a sensitivity of 0.95, specificity of 0.73, PPV of 0.56, and NPV of 0.97 were produced (Stone et al., 2008). The significantly higher rate of false positives was found to occur in children 12-13 months old. In removing these 12- and 13-month-olds from the sample, psychometrics improved; sensitivity was 0.93, specificity was 0.83, PPV was 0.68, and NPV was 0.97. Therefore, the *STAT* appears to be a promising tool for autism screening. However, more research should be conducted on larger samples (Campbell et al., 2016). Further, clinicians who use the *STAT* should be trained, should have time to accurately administer the measure, and should have experience in working with young children with ASD (Zwaigenbaum & Stone, 2006).

**Baby and Infant Screen for Children with aUtIsm Traits-Part 1 (BISCUIT-Part 1)**

The *BISCUIT* (Matson, Boisjoli, & Wilkins, 2007) is a three-part, parent report battery that assesses autism symptoms, comorbid psychopathology, and challenging behaviors in children aged 17-37 months old. The *BISCUIT-Part 1* measures ASD symptomology and is

made up of 62 items addressing social and communication skills and RRBIs. Caregivers are asked to rate each of the items on a 3-point scale comparing their child to children of the same age. To help informants rate each item, the *BISCUIT* includes an appendix that presents typical and atypical characteristics related to each item and concrete examples of relevant behaviors. With 62 items, the *BISCUIT-Part 1* is quite long in relation to other ASD screening tools available and can be time-intensive. As such, recent efforts have been made in regards to developing a briefer screener derived from the *BISCUIT* (Cervantes et al., in press; LoVullo & Matson, 2012). These developments will be discussed further within the Purpose and Method sections.

For the complete *BISCUIT-Part 1*, 276 children enrolled in an early intervention program who had a developmental delay or a condition likely to result in developmental delay were sampled for item selection and calculation of internal reliability. Of an initial pool of 71 items, 62 items were retained following examination of endorsement rates, corrected item-scale correlations, and inter-item correlations. Internal reliability of the final scale was estimated at 0.97 (Matson, Wilkins, Sevin, et al., 2009). In a follow-up study conducted with a sample of 1,007 children recruited from the same early intervention program, the authors of the scale sought to establish an optimal cutoff score and sensitivity and specificity values for the *BISCUIT-Part 1* (Matson, Wilkins, Sharp, et al., 2009). A cutoff of 17 on the complete measure corresponded with a sensitivity of 0.934 and a specificity of 0.866 when differentiating between ASD and no ASD. Further, an overall correct classification rate was found at 0.888 (Matson, Wilkins, Sharp, et al., 2009).

Later, factor analytic research revealed the *BISCUIT-Part 1* is composed of three distinct factors: socialization/non-verbal communication, RRBIs, and communication (Matson, Boisjoli,

Hess, & Wilkins, 2010); however, these results require replication with a *DSM-5* sample. Evidence of convergent validity has been shown through large positive correlations with the *M-CHAT* ($r$ = 0.80) and large negative correlations with Personal-Social domain scores of the *Battelle Developmental Inventory, Second Edition* (*BDI-2*; r = -0.50). Divergent validity of the *BISCUIT-Part 1* was established through small to moderate correlations with the Adaptive, Motor, and Cognitive domains of the *BDI-2* ($r$ = -0.19 for Adaptive, -0.30 for Motor, and -0.44 for Cognitive; Matson, Wilkins, & Fodstad, 2011).

More recent research has focused on developing distinct age-based cutoff scores for the *BISCUIT-Part 1* (Horovitz & Matson, 2013; Konst, Matson, & Matheis, submitted). Results from these studies have revealed higher cutoff scores for younger children are necessary. Also of note, research is now being conducted on the *BISCUIT* to account for the recent *DSM* changes. For example, in a study developing cutoff scores in a *DSM-5* diagnosed sample, estimates of sensitivity and specificity were found at 0.94 and 0.87, respectively (Konst et al., submitted). Though the *BISCUIT-Part 1* has been extensively studied with large samples and psychometrics appear promising, improvements on the measure's clinical utility in regards to time intensiveness may be beneficial.

# CHAPTER 6. HETEROGENEITY IN ASD

Autism as a spectrum disorder involves wide heterogeneity in the features of individuals diagnosed. Though all individuals with ASD present with significant social communication and interaction deficits and RRBIs, symptom presentation and severity, symptom onset patterns, etiological factors, and conditions that co-occur with autism can vary significantly across the ASD population (Johnson et al., 2007; Zwaigenbaum, 2010). This variability complicates ASD early detection and screening efforts (Goin-Kochel et al., 2006; Zwaigenbaum, 2010). Among the variables that may affect assessment accuracy and early ASD identification are child age, gender, and level of intellectual deficit. These variables are discussed further below and were examined in the current study.

## Age

Though diagnoses made as early as 2 years of age have been repeatedly shown stable through the lifespan in the research literature, change in symptom severity, symptom subtype, and co-occurring conditions across time has been noted (Cheslack-Postava & Jordan-Young, 2012; Guthrie et al., 2013; Matson & Nebel-Schwalm, 2007; McGovern & Sigman, 2005; L. M. Turner, 2006). In addition, the manner in which atypicalities are assessed is designed according to age-based behavioral expectations. For example, an infant with autism may be flagged for inconsistent responding to his/her name or lack of response to social stimuli. Concern for a toddler's development may result from delayed speech, lack of or limited functional and make-believe play, and RRBIs. An older child, adolescent, or adult may have difficulty forming and maintaining relationships, engaging in conversation, and/or understanding social subtleties (Cheslack-Postava & Jordan-Young, 2012).

In regards to symptom severity, researchers have found that many individuals with autism demonstrate improvements across the lifespan. Caregivers of individuals with ASD have reported fewer symptoms in adolescence and early adulthood compared to toddlerhood and early childhood (McGovern & Sigman, 2005). Improvements have been noted in cognitive and language scores from toddlerhood to middle childhood and in the areas of social interaction, RRBIs, adaptive behavior, and emotional responsiveness from childhood to adolescence; however, significantly more improvement in symptomology has been observed in individuals with high-functioning autism compared to those with low-functioning autism (McGovern & Sigman, 2005; L. M. Turner, 2006). Of note, many of these studies rely on the retrospective method of parent report; therefore, it is possible that caregivers are habituating to their child's behavior and rating current behavior as less severe than in the past (McGovern & Sigman, 2005).

Particularly relevant to the early identification of autism, Guthrie and colleagues (2013) examined change in symptology across infant and toddlerhood in individuals diagnosed with ASD. In this study, stability of diagnosis and symptoms were examined with 82 young children who underwent two comprehensive assessments at around 20 months old and again at 36 months old. Findings supported the stability of diagnoses across toddlerhood; all children diagnosed with ASD at the first assessment continued to meet criteria at the second assessment. Similarly, all children who were not given a diagnosis of ASD or whose diagnosis was not deferred did not meet criteria for ASD at the second assessment. While stability was shown in diagnosis, significant changes in symptom severity was observed over the course of 1-2 years in infant and toddlerhood. Infants and toddlers with ASD as well as those without ASD experienced improvement in social communication and interaction skills from the first to second evaluation. Children with ASD demonstrated an increase of RRBIs across infant and toddlerhood; this trend

was not seen in children without ASD. These symptom changes were associated with child developmental level at initial evaluation (Guthrie et al., 2013).

Also important in regards to infant and toddler development and ASD is the fact that a significant proportion of children diagnosed with autism experience a late symptom onset pattern (Johnson et al., 2007; Ozonoff et al., 2008). As previously mentioned, regression or developmental plateuing can occur before 3 years of age (Guthrie et al., 2013; Ozonoff et al., 2008). Further, typical development can be variable and early developmental concerns that may seem consistent with autism can resolve with time (Barton et al., 2012). These findings hold important implications for early identification and diagnosis of ASD. First, because regression is common in toddlerhood, routine ASD screening should be conducted at several time points as to ensure children who regress are detected (Johnson et al., 2007). Additionally, variation in typical development may lead to an increased rate of false positives on ASD screens early on. Screening data on very young children is limited and, more research is required to establish evidence of psychometric strength for the use of screening tools to detect risk in children under 2 years old (Barton et al., 2012). Beyond screening implications, age plays a role in the diagnosis of milder presentations of autism. Often, those children whose ASD symptoms are less severe are not diagnosed until later in life when social expectations increase and their impairments become more pronounced when compared to peers (Goin-Kochel et al., 2006).

## Gender

One of the most consistent findings in ASD research is the high male-to-female ratio. The accepted ratio for all ASD severity levels is approximately 4:1; however, the ratio increases to around 10:1 in cases of high-functioning autism and decreases to approximately 2:1 in cases of comorbid ASD and ID (Cheslack-Postava & Jordan-Young, 2012; Dworzynski, Ronald,

Bolton, & Happé, 2012). Therefore, many more females diagnosed with ASD are severely impaired with low intellectual ability and/or co-occurring behavior problems than are high-functioning (Dworzynski et al., 2012). Factors that have been proposed to account for this large discrepancy relate to both biological vulnerabilities and social, cultural, and environmental variables (Cheslack-Postava & Jordan-Young, 2012). Neurological and biological differences between males and females with ASD have been consistently found in the research literature leading to the theory that sex mediates the neurobiology of autism (Cheslack-Postava & Jordan-Young, 2012; Kirkovski, Enticott, Hughes, Rossell, & Fitzgerald, 2016; Kreiser & White, 2014; Rivet & Matson, 2011; Schaafsma & Pfaff, 2014). Social, cultural, and environmental variables may also relate to both symptom manifestation as well as interpretation of ASD symptoms (Hiller, Young, & Weber, 2015; Kreiser & White, 2014; Postorino et al., 2015). Kreiser and White (2014) noted several factors that may lead to differences in symptom presentation between genders. First, because gender role expectations for behavior emphasize the importance of empathy and social sensitivity in women, there may be more frequent or more potent consequences for engaging in disruptive or non-conforming behavior for females with ASD compared to males. This may influence the severity level of social deficits and associated externalizing problems. Further, children tend to associate with same-gender peers and engage in gender-typed play; female children are expected to engage socially with smaller, more intimate groups and in more interpersonally focused conversations. These expectations for behavior as well as the peer modeling that occurs when females with ASD associate with typically developing females may lead to improvements in conversation skills and emotional responsivity (Kreiser and White, 2014).

However, theses differences in ASD manifestation between genders have not been consistently found across studies (Postorino et al., 2015; Rivet & Matson, 2011). In fact, most findings show no differences in core ASD symptoms and ASD severity between males and females (Postorino et al., 2015). This lack of differentiation in ASD symptomology across genders has been evident in infants and toddlers as well (Sipes, Matson, Worley, & Kozlowski, 2011). Although much of the recent literature negates presentation differences between males and females, several studies have found disparities in the rate of RRBIs. Females with ASD have been found to engage in significantly less RRBIs than males (Frazier, Georgiades, Bishop, & Hardan, 2014; Hiller et al., 2015; Postorino et al., 2015; Rivet & Matson, 2011; Sipes et al., 2011; Van Wijngaarden-Cremers et al., 2014). However, this relationship may be moderated by developmental level. For example, Sipes and colleagues (2011) found that only female children with ASD and average developmental functioning demonstrated significantly fewer RRBIs, while females with ASD and low developmental functioning did not show any differences from males with ASD.

In regards to social-communication skills, some studies have indicated that females experience greater impairments (Frazier et al., 2014) while others have shown that females are less impaired in this domain (Hiller et al., 2015; Lai et al., 2011). These inconsistencies may relate to sampling differences and the potential effects of lower cognitive ability in females. Some studies have also reported that females with ASD are more likely to experience comorbid internalizing problems than males; however, findings regarding differences in co-occurring conditions have been mixed as well (Postorino et al., 2015).

Regardless of potential symptom differences, females experience a lack of ASD diagnosis, delayed ASD diagnosis, or misdiagnosis more frequently than males (Rivet & Matson,

2011). In fact, diagnosis of ASD has been shown to occur later in females even when symptom severity is held constant (Dworzynski et al., 2012; Postorino et al., 2015; Russell, Steer, & Golding, 2011). Further, females with low cognitive ability and additional behavior problems are more likely to be diagnosed with ASD; and, females with ASD and no cognitive impairments are at a heightened risk for delayed diagnosis or failure to be diagnosed (Begeer et al., 2012; Dworzynski et al., 2012; Hiller et al., 2015). Several researchers believe that, in addition to social and cultural influences, there may be potential biases toward males in current diagnostic criteria and/or current clinical practices (Dworzynski et al., 2012; Hiller et al., 2015; Kreiser & White, 2014; Postorino et al., 2015). If biases in diagnostic criteria truly exist, ASD assessment measures may not be as valid for use with females compared to males (Kreiser & White, 2014). Therefore, clinicians should be mindful of the issues present in the timely diagnosis of females with ASD and should consider all possible social and cultural influences when carefully evaluating the symptoms of female clients (Kreiser & White, 2014).

### Global Developmental Delay/Intellectual Disability

GDD is a diagnosis given to children under the age of 5 years old characterized by significant delays in two or more domains of development (i.e., motor skills, speech and language, social skills, activities of daily living, cognition; APA, 2013; McDonald, 2006; Tirosh & Jaffe, 2011). Because measures of intellectual functioning are unreliable and unstable in young children, GDD serves as a placeholder diagnosis until accurate assessment can be performed (Levy, 2011; Tirosh & Jaffe, 2011). A substantial number of children given the diagnosis of GDD go on to meet diagnostic criteria for ID later in life (Tirosh & Jaffe, 2011); therefore, GDD is often viewed as a precursor to intellectual impairment. A diagnosis of ID is characterized by persistent and signficant deficits in cognitive and adaptive skills; level of ID

ranges from mild to profound and is assigned based upon functioning in conceptual, social, and practical domains (APA, 2013). ASD is often associated with both ID and GDD (Tirosh & Jaffe, 2011). In fact, 50-70% of the autism population has ID, and up to 40% of the ID population meets criteria for ASD. The prevalence of ASD within individuals with ID increases as a function of the severity of intellectual deficit (DiGuiseppi et al., 2010; Matson & Shoemaker, 2009; McCarthy et al., 2010; Peters-Scheffer, Didden, & Lang, 2016). Because of the substantial overlap between ID and ASD, researchers believe there to be a shared neurobiological mechanism that explains etiology of both disorders (DiGuiseppi et al., 2010).

Symptom severity and presentation have been shown to be significantly different when ASD occurs with ID compared to when ASD occurs alone. Individuals with ASD and ID have greater social deficits and language problems, higher rates of RRBIs, and more challenging behaviors that are more likely to persist over time (Matson & Shoemaker, 2009; McCarthy et al., 2010; Peters-Scheffer et al., 2016). In addition, higher rates of comorbid psychopathologies and health conditions have been found when ASD and ID co-occur (Matson & Shoemaker, 2009). Evidence for symptom differences in toddlers with varying developmental functioning levels has also been noted; toddlers with ASD and low developmental functioning evince more severe ASD symptomology compared to toddlers with ASD and typical developmental functioning (Sipes et al., 2011). In addition to symptom severity, different subtypes of symptoms have been found for individuals with ASD with and without ID. In relation to RRBIs, researchers have differentiated between the types of behavior engaged in by high and low functioning individuals (Gabriels, Cuccaro, Hill, Ivers, & Goldson, 2005; Peters-Scheffer et al., 2016; M. Turner, 1999). High-level RRBIs are those most often found in individuals with ASD and typical intellectual functioning; these behaviors are more complex and include unusual attachment to objects,

restricted interests, and repetitive speech. Low-level RRBIs are found in individuals with ASD and low intellectual ability; they are less complex and consist of sensory stereotypies, repetitive motor movements, and self-injurious behaviors (Gabriels et al., 2005; M. Turner, 1999).

Overall, assessment for ASD in individuals with ID and assessment for ID in individuals with ASD is complicated; many of the clinical features associated with the two disorders overlap. Social deficits, communication impairments, and stereotyped behaviors are observed in both ASD and in ID (Peters-Scheffer et al., 2016). Therefore, to meet for a diagnosis of ASD, social interaction and social communication deficits must be in excess of that which would be accounted for by ID alone (DiGuiseppi et al., 2010). Because of the complicated nature of the ASD and ID overlap, it is imperative that assessment measures (particularly Level 2 screening tools) can reliably differentiate between symptoms derived from ID and symptoms related to ASD. Lastly, findings on early detection of ASD in young children with intellectual impairments have been inconsistent. Some studies show a comorbid cognitive impairment leads to earlier age of diagnosis, some show comorbid cognitive impairments are associated with later age of diagnosis, and some studies have failed to find an association between cognitive impairment and age of ASD diagnosis (Daniels & Mandell, 2013). Due to the complexity of assessment in individuals with intellectual impairments and the importance of early ASD detection for appropriate treatment, more research is needed on this topic.

# CHAPTER 7. PURPOSE

Given the wide heterogeneity within the autism population, it is essential that screening tools be able to detect risk in a variety of ASD phenotypes. Therefore, researchers must work towards improving the sensitivity of assessment measures, while preserving specificity, to account for the wide range of features across individuals with autism (Durkin et al., 2015; Williams, 2006). In addition to psychometric strength in use across the spectrum, screening tools must be affordable, convenient, efficient, and acceptable for the setting of interest (Arunyanart et al., 2012; Dosreis et al., 2006). Standardized screening tools administered by PCPs and early childhood service providers currently offer the best available method for early detection of ASD risk (Gray & Tonge, 2005; Robins et al., 2016); therefore, continued work to improve not only screening tools but also screening procedures would be beneficial.

In regards to establishing effective screening procedures, early intervention programs are particularly in need. This is specifically true for those programs serving children birth to 3 years old with a developmental delay or a condition likely to result in developmental delay under the Individuals with Disabilities Education Act, Part C. Populations served by such programs have been found to have a larger ASD prevalence compared to the general population. For example, researchers have shown that risk for autism is higher among individuals born preterm and/or with low birth weight, individuals with seizures, and individuals with genetic disorders like Down syndrome and Fragile X syndrome (Cervantes et al., in press; DiGuiseppi et al., 2010; Johnson et al., 2007; Lampi et al., 2012; Limperopoulos et al., 2008; Saemundsen, Ludvigsson, Hilmarsdottir, & Rafnsson, 2007). When children are enrolled in these early intervention programs, they are most likely receiving non-specific interventions that target more general deficits (Pizur-Barnekow, Muusz, McKenna, O'Connor, & Cutler, 2012). Given the large

evidence base for autism-specific interventions, early intervention providers and service coordinators play an integral role in identifying early symptoms of ASD and making referrals for more appropriate diagnostic and intervention services (Pizur-Barnekow et al., 2012). These factors combine to emphasize the importance of conducting routine autism screens in early intervention settings. Early intervention programs have several distinct features that may mediate some of the screening barriers described in the research in primary care settings. First, early intervention providers have very frequent contact with participants and their families; providers are frequently conducting therapy within participant homes and thus are readily available to field parental concerns. Next, early intervention providers are often trained and experienced in administering formal developmental measures. Therefore, these professionals may be more comfortable administering standardized ASD screening tools. Last, many early intervention providers have completed more autism-specific training compared to PCPs (Cervantes et al., in press).

As part of their autism initiative, EarlySteps, Louisiana's early intervention program under the Individuals with Disabilities Education Act, Part C, provides caregivers the opportunity for routine and formal autism screening of participants ages 1.5-3 years old. At intake, annual, and exit visits, caregivers are given the option to complete the *BISCUIT-Part 1*. Participants that score in the at-risk range on the measure are then referred to a specialty clinic for further evaluation. However, in line with the research findings within primary care settings (Crais et al., 2014; Gillis, 2009; Johnson et al., 2007), low rates of ASD screening have been reported within this program. Many families are declining the *BISCUIT-Part 1*. In fact, Matheis and Matson (2015) found that nearly a quarter of EarlySteps autism screens are declined. At higher likelihood of screen refusal were participants who were female, younger, and participants

with diagnoses of Down syndrome, cerebral palsy, seizure disorders, and other genetic/medical conditions. Caregivers of children with low birth weight, with a family history of autism, and African American children were less likely to refuse the ASD screen (Matheis & Matson, 2015).

In attempts to develop a more efficient screening strategy and thus encourage higher rates of accepted screens, LoVullo and Matson (2009) established the first abbreviated scoring algorithm for the *BISCUIT-Part 1*. This algorithm consisted of five items (i.e., [59] Development of social relationships, [53] Use of nonverbal communication, [4] Engages in repetitive motor movements for no reason, [17] Shares enjoyment, interests, or achievements with others, and [19] Interest in participating in social games, sports, and activities) that mapped onto the *BISCUIT* total score. Using a cutoff score of 2, sensitivity was found to be 0.941, specificity was 0.947, and PPV was 0.883 (LoVullo & Matson, 2012). However, LoVullo and Matson's (2009) algorithm was developed on a sample of approximately 2,000 EarlySteps participants according to the autism criteria within the *DSM-IV-TR* and was not cross-validated. Therefore, further research was necessary.

Cervantes and colleagues (in press) aimed to build upon the results of LoVullo and Matson's (2009) study. With a sample of 6,003 EarlySteps participants, the abbreviated scoring algorithm for the *BISCUIT-Part 1* was updated according to *DSM-5* ASD criteria. After performing cross-validation, a six-item solution proved optimal. Items included in the algorithm were: [8] Maintains eye contact, [53] Use of nonverbal communication, [10] Social interactions with others his/her age, [59] Development of social relationships, [34] Abnormal preoccupation with the parts of an object or objects, and [4] Engages in repetitive motor movements for no reason. Therefore, two social interaction and social communication items changed from the LoVullo and Matson (2009) algorithm; [19] Interest in participating in social games, sports, and

activities and [17] Shares enjoyment, interests, or achievements with others were replaced by [8] Maintains eye contact and [10] Social interactions with others his/her age. [34] Abnormal preoccupation with parts of an object or objects was added reflecting an increase in RRBI items represented within the algorithm. Establishing a cutoff score of 3 yielded a sensitivity of 0.960, a specificity of 0.864, a PPV of 0.502, and an NPV of 0.993. These psychometric estimates are similar to that of the complete *BISCUIT-Part 1* demonstrating that the algorithm can reliably distinguish children at-risk for ASD from those without risk (Cervantes et al., in press).

Using the abbreviated scoring algorithm as a screener may be advantageous locally for the EarlySteps program as well as globally for PCPs who are experiencing common barriers to screening. This may be particularly true when it is used within a two-pronged assessment approach. First and foremost, difficulties with time intensiveness may be remediated; offering sequential screening would be time efficient while remaining psychometrically sound (Campbell et al., 2016). This would involve first administering the algorithm to all participants; those who screen negative at this stage would avoid the lengthiness of the complete *BISCUIT-Part 1*. Those who screen positive would then be encouraged to complete the full 62-item measure; this would result in a decrease of false positives from the initial screen and prevent the anxiety and time needed to seek full comprehensive evaluation for those children who screened false positive on the algorithm. Finally, the participants who score at-risk on both the algorithm and the complete *BISCUIT-Part 1* would be referred to a specialty diagnostic clinic. Other advantages relate to the *BISCUIT*'s formatting and design. Because the *BISCUIT* is a parent report measure, the burden screening places on service providers may be reduced and accuracy may be improved as parents are more familiar with their child's behavior across settings and time (Barton et al., 2012; Robins et al., 2001). Additionally, researchers have indicated that many healthcare

providers reported concerns of parental difficulties completing ASD screening measures (Crais et al., 2014). The *BISCUIT* appendix may help to resolve the comprehension issues faced by caregivers.

Although use of the algorithm as a screener offers many benefits, more research is necessary to validate its clinical utility and psychometric properties across ASD subgroups. Research regarding potential effects of demographic and clinical factors has been limited within many available ASD screening tools; it is possible that existing algorithms may not be adequate for all presentations of autism and thus result in higher rates of inaccurate results (i.e., more false positive and false negative screens; DiGuiseppi et al., 2010). Therefore, the current paper aimed to assess the utility of the *BISCUIT* algorithm as a screener within young children of varying ages, genders, and levels of developmental functioning. The paper was divided into three distinct studies. In Study 1, psychometric properties of the abbreviated scoring algorithm were explored within children of three different age groups (i.e., 17-23 months old; 24-30 months old; 31-37 months old). Next, the performance of the algorithm with female versus male participants was compared in Study 2. In Study 3, psychometrics were examined for the use of the screener with children of low developmental functioning and children of typical developmental functioning. Given the low rate of completed screens within both the EarlySteps program and more universally in primary care settings, further development of the *BISCUIT-Part 1* abbreviated scoring algorithm could prove valuable.

# CHAPTER 8. GENERAL METHOD

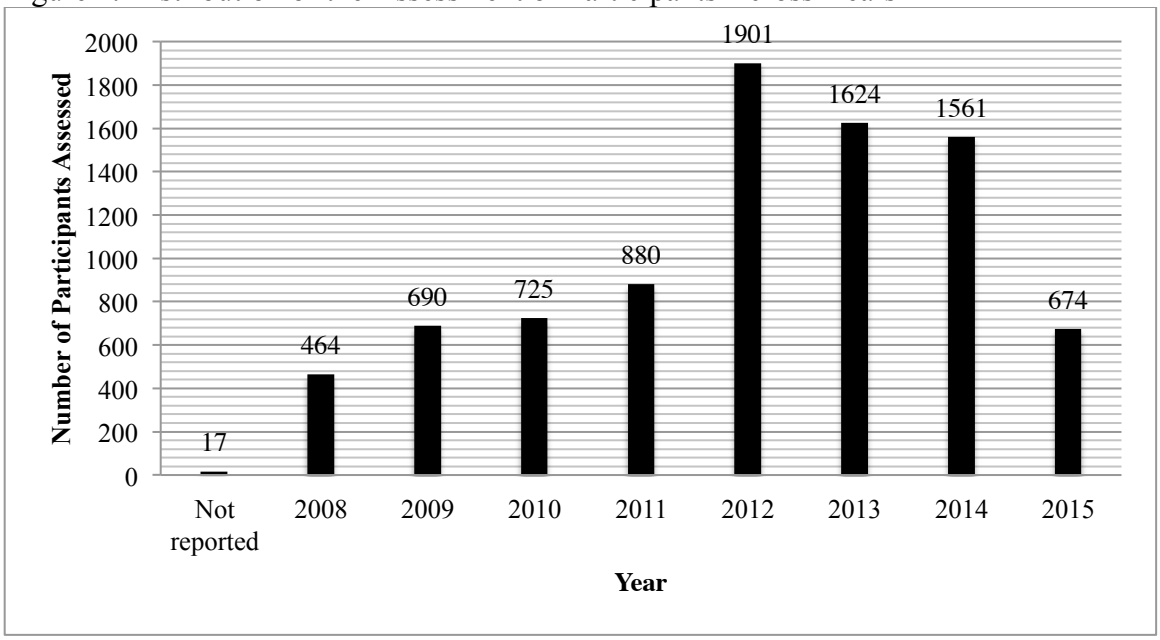
## Participants

The study's participants were pulled from a preexisting dataset and included infants and toddlers, ages 17-37 months, enrolled in EarlySteps. As previously mentioned, EarlySteps provides services to infants and toddlers from birth to 3 years old under the Individuals with Disabilities Education Act, Part C. The total sample size was 8,536 participants and reflects the combination of participants that were not missing relevant data points from both the LoVullo and Matson (2009) and the Cervantes et al. (in press) studies. This sample was divided into two groups based upon the presence of ASD. ASD diagnoses were assigned by a licensed doctoral level psychologist with over 30 years of experience and were based upon a *DSM-5* ASD criteria algorithm. All participants who were assigned an ASD diagnosis were placed in the ASD group. Participants who did not meet criteria for ASD were placed in an atypical development group; these children had a DD (e.g., cerebral palsy) or a condition likely to result in developmental delay (e.g., premature birth, low birth weight). The sample was broken down further to address each condition of interest according to several associated features (i.e., age, gender, level of developmental functioning). The procedures for grouping participants based upon associated features are discussed more comprehensively within each study's subsection. Demographic factors for the total sample are presented in Table 1. Of note, participants were assessed between the years of 2008 and 2015. Information regarding the proportion of children assessed between those years is available in Figure 1.

Table 1. Total Sample Demographics

| | Atypical Development (*n*=7397) | ASD (*n*=1139) | Total Sample (*n*=8536) |
|---|---|---|---|
| **Age (months)** | | | |
| ***M* (*SD*)** | 25.29 (4.62) | 26.07 (4.56) | 25.39 (4.62) |
| **Gender %(*n*)** | | | |
| Male | 67.53% (4995) | 76.21% (868) | 68.69% (5863) |
| Female | 32.47% (2402) | 23.79% (271) | 31.31% (2673) |
| **Ethnicity %(*n*)** | | | |
| African American | 36.50% (2700) | 41.70% (475) | 37.20% (3175) |
| Caucasian | 52.01% (3847) | 45.92% (523) | 51.20% (4370) |
| Hispanic | 3.62% (268) | 3.34% (38) | 3.58% (306) |
| Other/Unspecified | 7.87% (582) | 9.04% (103) | 8.02% (685) |

Figure 1. Distribution of the Assessment of Participants Across Years




## Measures

### Baby and Infant Screen for Children with aUtIsm Traits, Part 1 (BISCUIT – Part 1)

Previously discussed in the ASD screening tools subsection, the *BISCUIT-Part 1* is one measure within the three-part *BISCUIT* battery used to assess children aged 17-37 months. The *BISCUIT-Part 1* is designed to evaluate autism symptoms and includes 62 items rated on a 3-point Likert scale. Within the scale, a rating of "0" suggests "not different; no impairment"

compared to similarly aged children; a rating of "1" suggests "somewhat different; mild impairment;" and a rating of "2" suggests "very different; severe impairment." The ratings on the items summate to a total score. Based upon the total score, the child being assessed can fall into either an ASD range or an atypical development range. These cutoff ranges were developed on a sample of 6,860 infants and toddlers with and without *DSM-5* ASD diagnoses (Konst et al., submitted). Administration time needed for the complete *BISCUIT* is approximately 30 minutes.

Following concerns regarding EarlySteps' high rates of screen refusals, efforts towards creating an abbreviated scoring algorithm were made (Matheis & Matson, 2015; Cervantes et al., in press). Discussed previously, the most recent update to the algorithm consists of six items that were found to reliably differentiate children with ASD from those without. Four of the six items in the algorithm are related to social communication and social interaction; whereas, two items are related to RRBIs. Using a cutoff score of 3, sensitivity was found at 0.960, specificity at 0.864, PPV at 0.502, and NPV at 0.993 (Cervantes et al., in press).

As previously stated, the performance of the abbreviated scoring algorithm in its use with various demographic and clinical characteristics was explored in this study. However, if the items in the algorithm demonstrated inadequate psychometrics with any of the subgroups of interest, the complete *BISCUIT-Part 1* was to be utilized in developing potential revisions for its use with these populations.

**Battelle Developmental Inventory, Second Edition (BDI-2)**

The *BDI-2* (Newborg, 2005) measures the developmental functioning of children aged birth to 7 years, 11 months. Assessment across five domains (i.e., Adaptive, Personal-Social, Communication, Motor, and Cognitive) is conducted through both informant report and structured observation. The *BDI-2* consists of 450 items rated on a 3-point scale. Scores are

assigned either by a trained clinician or the child's caregivers; a score of "0" suggests "no ability in the skill," a score of "1" suggests "emerging ability," and a score of "2" suggests "ability in the skill." Item ratings are summed to calculate a total developmental quotient (DQ), DQs for each domain, and subdomain scores. Total DQ and the domain scores have a mean of 100 and a standard deviation of 15.

Evidence of adequate reliability and validity has been established for the *BDI-2*. Test-retest reliability has been found above 0.90 across total DQ and all domain scores. Internal reliability of the total DQ was estimated at 0.99 (Newborg, 2005). Convergent validity of the *BDI-2* was proven through correlations with scores from multiple well-established measures of child development (e.g., the *Bayley Scales of Infant Development, Second Edition [BSID-II],* the *Denver Developmental Screening Test, Second Edition [DDST-II*], the *Preschool Language Scales, Fourth Edition [PLS-4],* the *Wechsler Preschool and Primary Scale of Intelligence, Third Edition [WPPSI-III]*; Newborg, 2005). Total DQ was used within Study 3 for the assignment of participants into either low or typical developmental functioning subsamples.

## Procedure

All procedures were in accordance with ethical standards, and the study was approved by Louisiana's Office for Citizens with Developmental Disabilities (OCDD) and the Louisiana State University institutional review board. Prior to the receipt of the data used in this study, all personally identifying information about the participants was removed from the dataset. Because data was obtained from a deidentified database provided for research purposes, informed consent was not required from participants. Of note, informed consent was collected from all participants assessed prior to December 2013. Upon re-approval by the institutional review

boards in December 2013, it was determined that informed consent was not needed due to the deidentified nature of the database.

As part of intake, annual, and exit assessments, the *BISCUIT-Part 1* was offered to caregivers of children enrolled in EarlySteps as an optional ASD screener. Following caregiver agreement to screening, EarlySteps providers administered the *BISCUIT-Part 1* as part of a larger assessment battery. Within this battery, the *BDI-2* was also performed with EarlySteps participants at each intake, annual, and exit visit. The approximately 175 providers administering these measures held an appropriate degree as well as a certification or licensure in relevant fields (e.g., occupational therapy, physical therapy, speech and language pathology, psychology, social work, special education). The providers were trained on the administration of the *BISCUIT* and *BDI-2* and were experienced in evaluating and treating infants and toddlers. Of note, caregiver refusal of the ASD screen does not impact their child's qualification for early intervention services.

# CHAPTER 9. STUDY 1: AGE

ASD symptom emergence occurs over the first several years of life, and symptoms become more easily distinguishable with age (Goin-Kochel et al., 2006; Werner et al., 2005; Wetherby et al., 2006; Zwaigenbaum, 2010). As such, caregivers of younger children may be less likely to endorse impairment when assessing autism symptoms. On the contrary, four of the six items on the abbreviated scoring algorithm focus on behaviors related to social communication and social interaction. These behaviors typically develop over the course of the second year of a child's life, and variation in typical development is common (Baird et al., 2001; Stone et al., 2008). As such, caregivers of younger children may be more likely to endorse impairments on these items if the skills have not yet fully developed. Further, previous researchers have found that higher cutoff scores were necessary for younger samples on other screening tools and on the complete *BISCUIT-Part 1* (Konst et al., submitted; Stone et al., 2008).

Therefore, psychometric properties of the algorithm in its current form may be less adequate for younger samples. A new and likely higher cutoff score or perhaps the addition of more items could be necessary to improve the utility of the algorithm as a screener for this group. In order to examine the potential effects of age on the performance of the algorithm, the total sample was divided into three age groups (i.e., 17-23 months; 24-30 months; 31-37 months). These age groups were selected according to approximations for when developmental milestones are typically reached (e.g., first words, play skills, etc.) and are consistent with previous research (Horovitz & Matson, 2013; Konst et al., submitted). Each group consisted of participants with ASD and participants with atypical development.

## Study 1A: 17-23 Month Sample

### Method

**Participants.** The 17-23 month sample included 2969 participants; 332 were assigned to the ASD group and 2637 had atypical development. For purposes of examining psychometric properties as well as suggesting potentially necessary revisions, the 17-23 month sample was divided in half. A total of 1482 participants were placed in an exploratory group, and 1487 participants were placed in a revisions/replication group using random selection. This was done to ensure that validation could be performed with a separate subsample if the current algorithm resulted in inadequate psychometrics that required resolution (i.e., by creating an alternate algorithm with a different set of *BISCUIT* items). The exploratory and revisions/replication group did not differ in terms of gender ($X^2$ [1] = 1.23, $p$ = 0.28), ethnicity ($X^2$ [3] = 3.74, $p$ = 0.29), age ($F$ [1, 2697] = 1.49, $p$ = 0.22), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 0.19, $p$ = 0.68).

The exploratory group consisted of 162 participants who had ASD and 1320 participants who had atypical development. Within the exploratory group, no significant differences were found between participants with ASD and participants with atypical development in age ($F$ [1, 1481] = 0.15, $p$ = 0.70), gender ($X^2$ [1] = 0.78, $p$ = 0.38), or ethnicity ($X^2$ [3] = 4.43, $p$ = 0.22). The revisions/replication group included 170 participants who had ASD and 1317 participants who had atypical development. No significant age differences ($F$ [1, 1486] = 1.24, $p$ = 0.27) were found between participants with ASD and participants with atypical development. Significant differences were found in gender ($X^2$ [1] = 9.25, $p$ = 0.002) and ethnicity ($X^2$ [3] =8.45, $p$ = 0.04). However, gender differences were expected given the high male-to-female ratio in ASD (CDC, 2014); and, ethnicity differences were small and were not expected to

influence results. Further, symptom differences across ethnicities that have been noted are most likely attributable to disparity in ASD identification services and differences in cultural appraisal of typical development rather than true presentation differences (Mandell et al., 2009; Matson, Worley, et al., 2011). Demographic information for the exploratory and revisions/replication groups is presented in Table 2.

Table 2. 17-23 Month Sample Demographics

| | ***Exploratory Group*** | | ***Revisions/Replication Group*** | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 1320) | ASD ($n$ = 162) | Atypical Development ($n$ = 1317) | ASD ($n$ = 170) |
| **Age (months)** | | | | |
| *M* (*SD*) | 20.34 (1.83) | 20.40 (1.83) | 20.24 (1.78) | 20.41 (1.82) |
| **Gender %(*n*)** | | | | |
| Male | 64.39%(850) | 67.90%(110) | 65.38%(861) | 77.06%(131) |
| Female | 35.61%(470) | 32.10%(52) | 34.62%(456) | 22.94%(39) |
| **Ethnicity %(*n*)** | | | | |
| African American | 32.20%(425) | 38.27%(62) | 34.55%(455) | 43.53%(74) |
| Caucasian | 56.97%(752) | 49.38% (80) | 53.99%(711) | 42.35% (72) |
| Hispanic | 3.18%(42) | 4.94%(8) | 3.42%(45) | 3.53%(6) |
| Other/Unspecified | 7.65%(101) | 7.41% (12) | 8.05%(106) | 10.59%(18) |

**Research Design**. Within the exploratory group, logistic regression and receiver operating characteristic (ROC) analysis were used to examine model fit, Area Under the Curve (AUC) values, sensitivity and specificity estimates, and PPV and NPV for the current algorithm. The revisions/replication group was then to be used for two different functions based upon the results of these analyses in the exploratory group. If sensitivity and specificity estimates were found to be acceptable (above 0.80) in the exploratory group, the revisions/replication group was used to further validate the performance of the algorithm in predicting ASD risk. Therefore, the above statistics were repeated with this group. Of note, if the ROC curve demonstrated a different optimal cutoff score for discriminating between ASD and atypical development within these groups, age-based cutoff revisions were considered.

However, if sensitivity or specificity estimates were found to be below 0.80 in the exploratory group even with a change in cutoff score, the exploratory group was to be used to determine if a better set of items from the complete *BISCUIT-Part 1* exists to distinguish ASD from no ASD. To do this, a logistic regression was to be conducted to identify how well each of the 62 *BISCUIT-Part 1* items predicts the presence of ASD in 17- to 23-month-olds. Then, the ten items with the greatest odds ratios were further examined with a ROC analysis. The ROC analysis was used to determine the number of items needed to create an alternate algorithm with adequate sensitivity and specificity. As in the development of the current algorithm and due to the nature of diagnostic screeners, more weight was placed on sensitivity than specificity. The Area Under the Curve (AUC) values and the sensitivity and specificity tradeoffs for each potential algorithm were examined to identify the items to be retained. Following selection of an alternate algorithm, logistic regression and ROC analysis were conducted with the revisions/replication group to identify model fit and the cutoff score resulting in the best sensitivity and specificity estimates. Based upon the resulting cutoff score, PPV and NPV were calculated for the revisions/replication group.

**Results**

**Exploratory Group.** The assumptions of logistic regression were examined prior to running the analyses. First, the sample size far exceeded previous recommendations for the number of observations per predictor in a logistic regression (Peng, Lee, & Ingersoll, 2002). Further, examination of casewise diagnostics, including Cook's distance and leverage values, was used to detect outliers demonstrating undue influence; however, no such cases were identified. To examine the assumption of multicollinearity, tolerance and variable inflation factors (VIFs) were computed. For the six items in the algorithm, tolerance was greater than

0.10 (all range from 0.522-0.795) and VIFs were less than 10 (all range from 1.257-1.924) indicating that this assumption was met (Field, 2009). The logistic regression was then performed on the exploratory group to examine model fit; the model was statistically significant, $\chi^2$ (6) = 565.08, $p < .001$. All items of the algorithm were found to be significant predictors. The algorithm explained approximately 63.6% (Nagelkerke $R^2$) of the variance in presence of ASD and had an overall correct classification percentage of 94.1% (Table 3).

Table 3. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 1291 | 29 | 97.8 |
| **ASD** | 58 | 104 | 64.2 |
| **Overall** | - | - | 94.1 |

A ROC analysis was then conducted. The abbreviated scoring algorithm had an AUC value of 0.955 indicating excellent discriminating ability within this group (Compton, Fuchs, Fuchs, & Bryant, 2006). The ROC curve, shown in Figure 2, was then used to assess an optimal cutoff score for the algorithm. Identical to that found in Cervantes et al. (in press), the optimal cutoff score for this younger group was 3. This score yielded a sensitivity of 0.963 and a specificity of 0.851 (see Table 4). PPV and NPV were then computed in the exploratory group. PPV was 0.442, and NPV was 0.995 (presented in Table 5).

Figure 2. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

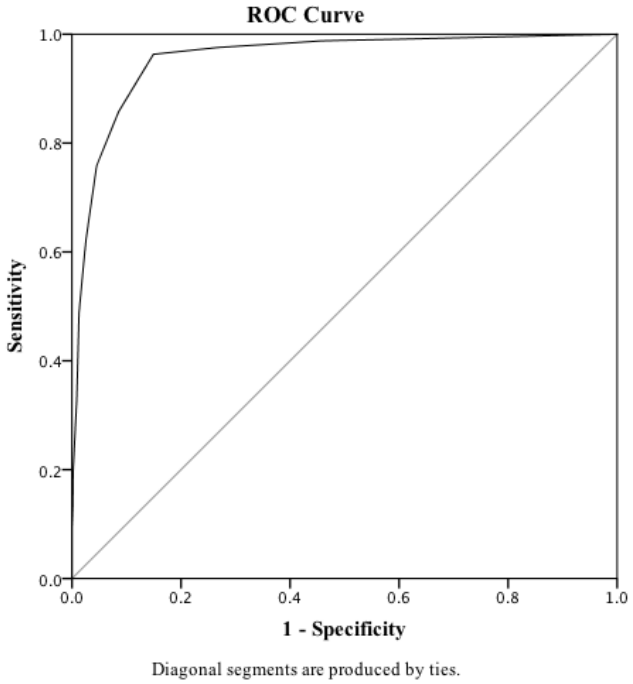



Table 4. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 0.988 | 0.544 |
| 2 | 0.975 | 0.736 |
| **3** | **0.963** | **0.851** |
| 4 | 0.858 | 0.914 |
| 5 | 0.759 | 0.955 |
| 6 | 0.623 | 0.974 |
| 7 | 0.488 | 0.987 |
| 8 | 0.321 | 0.992 |
| 9 | 0.198 | 0.998 |
| 10 | 0.136 | 0.998 |
| 11 | 0.093 | 1.000 |
| 12 | 0.056 | 1.000 |

Table 5. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 156 | 197 | 353 |
| **Negative Score (*n*)** | 6 | 1123 | 1129 |
| **Total (*n*)** | 162 | 1320 | 1482 |

**Revisions/Replication Group.** Prior to running analyses, the assumptions of logistic regression were checked using the same methods as in the exploratory group. Then to replicate the findings from the exploratory group, a logistic regression was performed on the revisions/replication group to examine model fit. As in the exploratory group, the regression model was statistically significant, $\chi^2$ (6) = 598.45, $p < .001$. However, all items of the algorithm but one (item [10] Social interactions with others his/her age) were found to be significant predictors. The algorithm explained approximately 65.1% (Nagelkerke $R^2$) of the variance in the presence of ASD and had an overall correct classification percentage of 93.1% (see Table 6).

Table 6. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 1282 | 35 | 97.3 |
| **ASD** | 68 | 102 | 60.0 |
| **Overall** | - | - | 93.1 |

The ROC analysis conducted after demonstrated that the algorithm had an AUC value of 0.948; this again indicated that the model had excellent discriminating ability within the revisions/replication group (Compton et al., 2006). The ROC curve (Figure 3) was then used to assess the optimal cutoff score found within the exploratory group. The cutoff score of 3 yielded a sensitivity of 0.918 and a specificity of 0.854. PPV and NPV estimates were 0.448 and 0.989, respectively (presented in Table 7).

Figure 3. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

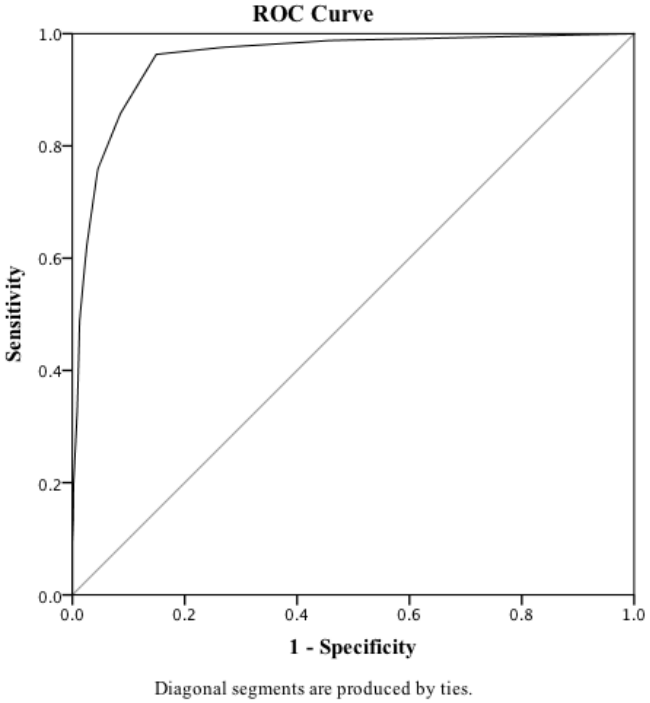



Table 7. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 156 | 192 | 348 |
| **Negative Score (*n*)** | 14 | 1125 | 1139 |
| **Total (*n*)** | 170 | 1317 | 1487 |

**Discussion**

Contrary to the hypothesis, no changes were necessary to the items of the abbreviated scoring algorithm or to the cutoff score for this younger sample. Similar to the original estimate of 0.960 from Cervantes et al. (in press), sensitivity of the algorithm for use with children 17-23 months old ranged from 0.918-0.963 in the current study. Specificity estimates were slightly lower, ranging from 0.851-0.854 compared to 0.864. PPV was again slightly lower at 0.442 in the exploratory group and 0.448 in the revisions/replication group compared to 0.502 in the Cervantes et al. (in press) study. NPV estimates were comparable in the 17-23 month old sample (NPV = 0.989-0.995) and the Cervantes et al. (in press) study (NPV = 0.993).

Table 8 displays the distribution of total scores across participants in both the exploratory and revisions/replication groups who scored true and false positive and true and false negative on the algorithm. Of the individuals who scored true positive, total scores varied widely across the group; however, the highest frequency score was 6. Compared to only 20.51% of the true positive group, nearly 70% of individuals scoring false positive had a total score of 3 or 4 (i.e., the lowest scores that could result in a positive screen). Of participants who scored true negative on the algorithm, a majority had a total score of 0. Compared to only 14.68% of the true negative group, 55% of the participants with ASD who scored negative on the algorithm had a total score of 2 (i.e., the highest score that could result in a negative screen).

Table 8. Total score on algorithm across participants classified both correctly and incorrectly

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | 26 | 28 | 47 | 49 | 44 | 32 | 29 | 19 | 12 | 16 | 312 |
| **True Negative** | 1422 | 496 | 330 | - | - | - | - | - | - | - | - | - | - | 2248 |
| **False Positive** | - | - | - | 157 | 112 | 58 | 32 | 15 | 11 | 2 | 2 | 0 | 0 | 389 |
| **False Negative** | 5 | 4 | 11 | - | - | - | - | - | - | - | - | - | - | 20 |

In the true positive group, the most frequent ratings of "2" were on items [10] Social interactions with others his/her age and [59] Development of social relationships. This highlights the significant emphasis on social impairment in the diagnosis of ASD. The true positive group was least likely to endorse item [34] Abnormal preoccupation with the parts of object(s); nearly half of the group received a rating a "0" on this item. Of note, item [34] was the lowest endorsed item across all participants in this sample. When participants scoring false positive endorsed symptoms, they were more likely to receive a rating of "1" than "2." The only item where this was not true was item [4] Engages in repetitive motor movements for no reason (28.53% were rated a "2;" 20.31% were rated a "1"). It is important to note that repetitive motor movements are not exclusive to ASD and should not be viewed as such in the diagnosis of DDs. Of participants who scored true negative, a rating of "0" was endorsed across all items for nearly

90% of participants in this group. For participants scoring false negative, the most frequently endorsed items were item [10] Social interactions with others his/her age, [59] Development of social relationships, and [8] Maintains eye contact. This trend again indicates that the social components of ASD are most important in diagnosis. Interestingly, item [10] Social interactions with others his/her age did not serve as a significant predictor in the revisions/replication group; this is likely due to the high rates of endorsement of impairment in this area across all groups. Therefore, peer interaction may be indicative of ASD and of more general developmental deficits in very young childhood.

## Study 1B: 24-30 Month Sample

### Method

**Participants.** The 24-30 month sample consisted of 4218 participants; 595 were in the ASD group and 3623 were in the atypical development group. As in the 17-23 month sample, the 24-30 month sample was divided in half to ensure that validation could be performed with a separate subsample if the current algorithm required revisions. A total of 2099 participants were placed in the exploratory group, and 2119 participants were assigned to the revisions/replication group using random selection. The exploratory and revisions/replication group did not differ in terms of gender ($X^2$ [1] = 0.86, $p$ = 0.35), ethnicity ($X^2$ [3] = 0.376, $p$ = 0.95), age ($F$ [1, 2697] = 1.49, $p$ = 0.22), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 3.16, $p$ = 0.08).

The exploratory group included 276 participants with ASD and 1823 atypically developing participants. Within the exploratory group, no significant differences were found between participants with ASD and participants with atypical development in ethnicity ($X^2$ [3] = 1.51, $p$ = 0.68). Significant differences between participants were found in regards to gender ($X^2$

[1] = 7.42, $p = 0.006$) and age ($F$ [1, 2097] = 5.60, $p = 0.02$, *partial* $\eta^2 = 0.003$). However, age differences were small, and gender differences again are to be expected given the nature of ASD. Within the revisions/replication group, 1800 participants had atypical development and 319 were assigned to the ASD group. No significant ethnicity differences were found between participants with ASD and participants with atypical development ($X^2$ [3] = 4.35, $p = 0.23$). However, similar to the exploratory group, significant differences were found in gender ($X^2$ [1] = 11.17, $p = 0.001$) and age ($F$ [1, 2117] = 8.32, $p = 0.004$, *partial* $\eta^2 = 0.004$). Again, gender differences were expected, and age differences were small. Demographic information for the exploratory group and the revisions/replication group is presented in Table 9.

Table 9. 24-30 Month Sample Demographics

| | ***Exploratory Group*** | | ***Revisions/Replication Group*** | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 1823) | ASD ($n$ = 276) | Atypical Development ($n$ = 1800) | ASD ($n$ = 319) |
| **Age (months)** | | | | |
| *M* (*SD*) | 26.59 (1.90) | 26.88 (1.98) | 26.65 (2.00) | 27.00 (1.96) |
| **Gender %(*n*)** | | | | |
| Male | 68.73% (1253) | 76.81% (212) | 67.06% (1207) | 76.49% (244) |
| Female | 31.27% (570) | 23.19% (64) | 32.94% (593) | 23.51% (75) |
| **Ethnicity %(*n*)** | | | | |
| African American | 37.69% (687) | 40.58% (112) | 38.00% (684) | 43.89% (140) |
| Caucasian | 50.96% (929) | 48.19% (133) | 50.50% (909) | 45.45% (145) |
| Hispanic | 3.73% (68) | 2.90% (8) | 3.67% (66) | 2.82% (9) |
| Other/Unspecified | 7.62% (139) | 8.33% (23) | 7.83% (171) | 7.84% (25) |

**Research Design.** The statistics used within the 17-23 month sample were repeated within the 24-30 month sample.

**Results**

**Exploratory Group.** Using the same methods as in Study 1A, the assumptions of logistic regression were examined and no issues arose. The logistic regression was then performed. The model was statistically significant, $\chi^2$ (6) = 950.77, $p < .001$, and all items of the

algorithm were found to be significant predictors. The model explained 67.3% (Nagelkerke $R^2$) of the variance in presence of ASD and had an overall correct classification percentage of 92.9% (see Table 10).

Table 10. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 1774 | 49 | 97.3 |
| **ASD** | 100 | 176 | 63.8 |
| **Overall** | - | - | 92.9 |

The subsequent ROC analysis indicated that the algorithm had an AUC value of 0.962 demonstrating excellent discriminating ability within this group (Compton et al., 2006). As in previous analyses, the ROC curve (Figure 4) was used to identify an optimal cutoff score within this subgroup. The cutoff score identified was 3 and yielded a sensitivity of 0.964 and a specificity of 0.843 (see Table 11). Within this exploratory group, PPV was 0.482 and NPV was 0.994 (see Table 12).

Figure 4. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

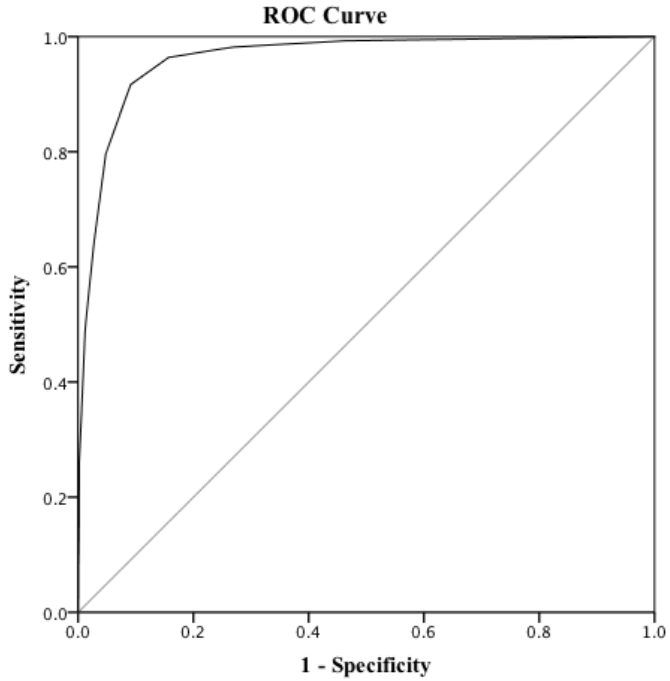


Table 11. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 0.993 | 0.536 |
| 2 | 0.982 | 0.731 |
| **3** | **0.964** | **0.843** |
| 4 | 0.917 | 0.909 |
| 5 | 0.797 | 0.952 |
| 6 | 0.638 | 0.973 |
| 7 | 0.489 | 0.988 |
| 8 | 0.362 | 0.993 |
| 9 | 0.261 | 0.998 |
| 10 | 0.159 | 0.998 |
| 11 | 0.083 | 0.999 |
| 12 | 0.043 | 0.999 |

Table 12. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 266 | 286 | 552 |
| **Negative Score (*n*)** | 10 | 1537 | 1547 |
| **Total (*n*)** | 276 | 1823 | 2099 |

**Revisions/Replication Group.** The assumptions of logistic regression were again checked prior to running replication analyses. The logistic regression was conducted, and the model was statistically significant, $\chi^2$ (6) = 1026.27, $p < .001$. All six items were found to be significant predictors. The model explained 67.2% (Nagelkerke $R^2$) of the variance in presence of ASD. An overall correct classification percentage of 92.5% was found (see Table 13).

Table 13. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 1749 | 51 | 97.2 |
| **ASD** | 107 | 212 | 66.5 |
| **Overall** | - | - | 92.5 |

As in the exploratory group, a ROC analysis was then conducted. The algorithm had an AUC value of 0.959 (i.e., excellent discriminating ability within this group; Compton et al., 2006). The ROC curve, presented in Figure 5, was then used to assess the optimal cutoff score

found in the exploratory group. The cutoff score of 3 resulted in a sensitivity of 0.950 and a specificity of 0.851. PPV was 0.531 and NPV was 0.990 (see Table 14).

Figure 5. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

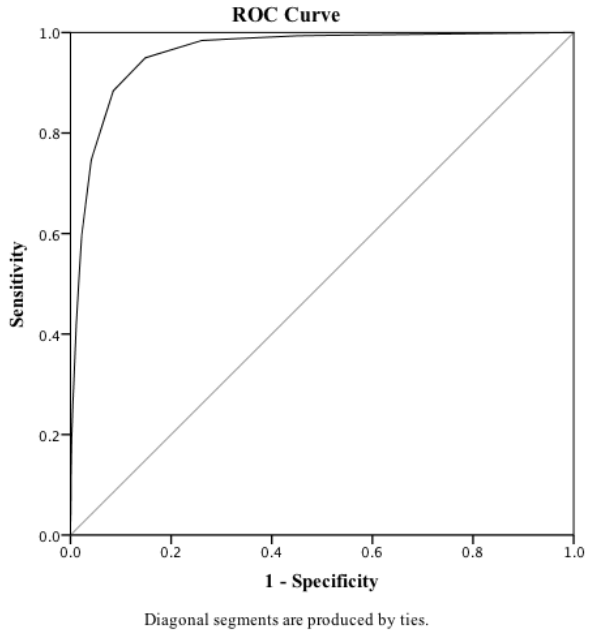



Table 14. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 303 | 268 | 571 |
| **Negative Score (*n*)** | 16 | 1532 | 1548 |
| **Total (*n*)** | 319 | 1800 | 2119 |

**Discussion**

As in the 17-23 month study, no changes were found necessary to the item make-up or to the cutoff score of the abbreviated scoring algorithm with use for children 24-30 months old. In regards to psychometric properties, the sensitivity within this sample ranged from 0.950-0.964 and was similar to that found in the Cervantes et al. (in press) and Study 1A: 17-23 Month Sample. Specificity was similar as well, ranging from 0.843-0.854. PPV estimates between 0.482-0.531 in the 24-30 month sample were slightly higher than PPV estimates in the 17-23 month sample (0.442-0.448) and more similar to the Cervantes et al. (in press) estimate

(PPV=0.502). NPV estimates across studies were also comparable (i.e., 0.993 in Cervantes et al. [in press]; 0.989-0.995 in the Study 1A: 17-23 Month; 0.990-0.994 in the current study).

The distribution of total scores across participants who scored true and false positives and true and false negative on the algorithm is presented in Table 15. The most frequent total score of children who scored true positive was a 6. Of the true negative group, 64.03% scored a 0 on the total algorithm, and 86.74% scored either a 0 or a 1. Nearly half of the children who scored false positive on the algorithm had a total score of 3, and over 70% of this group scored a 3 or a 4. Similarly, over 60% of the children who scored false negative on the abbreviated algorithm had a total score of 2.

Table 15. Total score on algorithm across participants classified both correctly and incorrectly

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | 34 | 77 | 92 | 93 | 79 | 61 | 54 | 41 | 19 | 19 | 569 |
| **True Negative** | 1965 | 697 | 407 | - | - | - | - | - | - | - | - | - | - | 3069 |
| **False Positive** | - | - | - | 235 | 157 | 73 | 45 | 21 | 15 | 3 | 3 | 1 | 1 | 554 |
| **False Negative** | 4 | 6 | 16 | - | - | - | - | - | - | - | - | - | - | 26 |

Of note, children who scored false positive were more likely to receive ratings of "1" than ratings of "2" across most items when impairment was endorsed. This was not true for those who scored true positive. The most frequent ratings of "2" in the true positive group were on item [10] Social interactions with others his/her age and item [59] Development of social relationships. Similarly, children who scored false negative were most likely to endorse at least some difficulties with item [10] Social interactions with others his/her age (42.31%) and item [8] Maintains eye contact (38.46%). As previously mentioned in the 17-23 month study, these results are not surprising given that social skill impairment is at the core of ASD. Relatedly, item [10] Social interactions with others his/her age was the most frequently endorsed item of the algorithm across the total sample. Over 50% of the true positive group were rated a "2" on this item and, nearly 95% of this true positive group endorsed some level of impairment. In contrast,

32.85% of the false negative group were rated a "2" on item [10] and, about 70% of this group endorsed some level of impairment on this item. Even within the true negative group, item [10] had the lowest frequency ratings of "0". Therefore, as seen also in the 17-23 month sample, problems with peer relations may be more universal within participants with ASD but difficulties in peer interactions are not exclusive to ASD.

In regards to RRBIs, item [34] Abnormal preoccupation with the parts of object(s) was the lowest rated item across groups. This result was consistent with the 17-23 month study. Also in line with the 17-23 month study, item [4] Engages in repetitive motor movements for no reason was the only item that the false positive group had more frequent ratings of "2" (26.35% of group) than "1" (20.04% of group). This again reiterates the importance of viewing repetitive motor movements as common in ASD but not exclusive to ASD. Further, nearly 75% of the false negative group did not endorse any impairment on item [4] Engages in repetitive motor movements for no reason demonstrating that not only are repetitive motor movements not exclusive to ASD but are also not necessary for an ASD diagnosis.

## Study 1C: 31-37 Month Sample

### Method

**Participants.** The 31-37 month old sample consisted of 1349 participants; 212 had ASD and 1137 had atypical development. As in previous analyses, the sample was divided in half. Using random selection, a total of 652 participants were placed in the exploratory group, and 697 participants were placed in the revisions/replication group. The exploratory and revisions/replication group did not differ in terms of gender ($X^2$ [1] = 0.01, $p$ = 0.91), ethnicity ($X^2$ [3] = 6.75, $p$ = 0.08), age ($F$ [1, 1347] = 1.75, $p$ = 0.19), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 1.25, $p$ = 0.26).

The exploratory group consisted of 95 participants with ASD and 557 with atypical development. No significant differences were found between participants with ASD and participants with atypical development in age ($F$ [1, 650] = 0.26, $p$ = 0.61) and ethnicity ($X^2$ [3] = 6.64, $p$ = 0.08). As in previous analyses, significant gender differences were found ($X^2$ [1] = 4.12, $p$ = 0.04). In the revisions/replication group, 580 participants had atypical development and 117 had ASD. No significant differences were found between participants with ASD and with atypical development in gender ($X^2$ [1] = 2.29, $p$ = 0.13), ethnicity ($X^2$ [3] = 0.89, $p$ = 0.83), or age ($F$ [1, 695] = 1.83, $p$ = 0.12) within this group. Demographic information for the exploratory and the revisions/replication group is presented in Table 16.

Table 16. 31-37 Month Sample Demographics

| | ***Exploratory Group*** | | ***Revisions/Replication Group*** | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 557) | ASD ($n$ = 95) | Atypical Development ($n$ = 580) | ASD ($n$ = 117) |
| **Age (months)** | | | | |
| *M* (*SD*) | 32.56 (1.34) | 32.48 (1.37) | 32.68 (1.44) | 32.49 (1.29) |
| **Gender %(*n*)** | | | | |
| Male | 72.17% (402) | 82.11% (78) | 72.76% (422) | 79.49% (93) |
| Female | 27.83% (155) | 17.89% (17) | 27.24% (158) | 20.51% (24) |
| **Ethnicity %(*n*)** | | | | |
| African American | 43.45% (242) | 42.11% (40) | 35.70% (207) | 40.17% (47) |
| Caucasian | 45.60% (254) | 41.05% (39) | 50.34% (292) | 46.15% (54) |
| Hispanic | 3.77% (21) | 2.10% (2) | 4.48% (26) | 4.28% (5) |
| Other/Unspecified | 7.18% (40) | 14.74% (14) | 9.48% (55) | 9.40% (11) |

**Research Design.** The statistics used within the 17-23 month sample were repeated within the 31-37 month sample.

**Results**

**Exploratory Group.** Assumptions were checked prior to running analyses. The logistic regression was performed, and the regression model was statistically significant, $\chi^2$ (6) = 313.96, $p$ < .001. All six items were significant predictors. The model explained 67.7% (Nagelkerke $R^2$)

of the variance in presence of ASD. The overall correct classification percentage was 92.2% (see Table 17).

Table 17. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 539 | 18 | 96.8 |
| **ASD** | 33 | 62 | 65.3 |
| **Overall** | - | - | 92.2 |

A ROC analysis was then performed. The algorithm had an AUC value of 0.961 indicating excellent discriminating ability within this group (Compton et al., 2006). Utilizing the ROC curve shown in Figure 6, a cutoff score of 3 was identified as optimal. This score yielded a sensitivity of 0.958 and a specificity of 0.853 (see Table 18). Within this exploratory group, PPV was estimated at 0.526 and NPV was 0.992 (see Table 19).

Figure 6. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

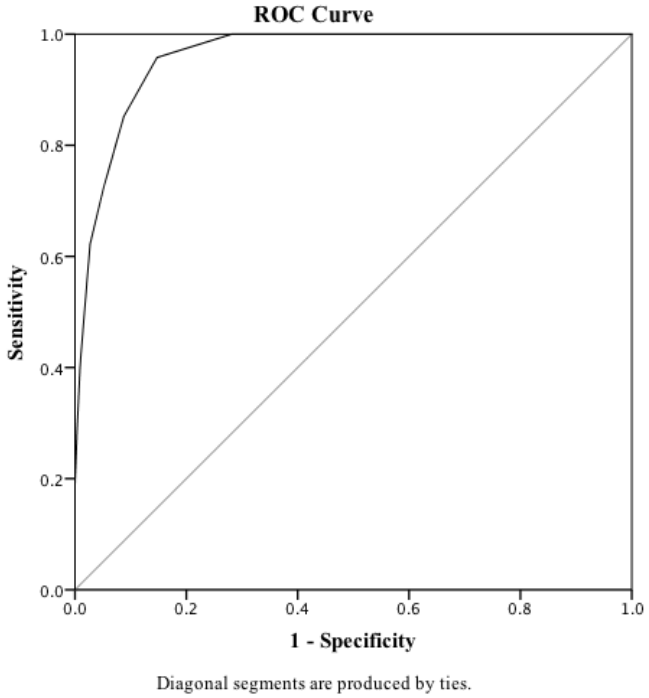


Table 18. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 1.000 | 0.533 |
| 2 | 1.000 | 0.718 |
| **3** | **0.958** | **0.853** |
| 4 | 0.853 | 0.912 |
| 5 | 0.726 | 0.948 |
| 6 | 0.621 | 0.973 |
| 7 | 0.400 | 0.991 |
| 8 | 0.284 | 0.996 |
| 9 | 0.189 | 1.000 |
| 10 | 0.126 | 1.000 |
| 11 | 0.063 | 1.000 |
| 12 | 0.032 | 1.000 |

Table 19. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 91 | 82 | 173 |
| **Negative Score (*n*)** | 4 | 475 | 479 |
| **Total (*n*)** | 95 | 557 | 652 |

**Revisions/Replication Group.** Assumptions were checked, and the logistic regression was then performed. The model was statistically significant, $\chi^2$ (6) = 371.35, $p < .001$. All six items were found to be significant predictors. The model explained 69.4% (Nagelkerke $R^2$) of the variance in presence of ASD within the revisions/replication group and had an overall correct classification percentage of 92.1% (see Table 20).

Table 20. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 560 | 20 | 96.6 |
| **ASD** | 35 | 82 | 70.1 |
| **Overall** | - | - | 92.1 |

The subsequent ROC analysis resulted in an AUC value of 0.953 (i.e., excellent discriminating ability within this group; Compton et al., 2006). The ROC curve (Figure 7) was then used to assess the optimal cutoff score of 3 found in the exploratory group. A sensitivity of

0.923 and a specificity of 0.857 were calculated. In the revisions/replication group, PPV was 0.565 and NPV was 0.982 (see Table 21).

Figure 7. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

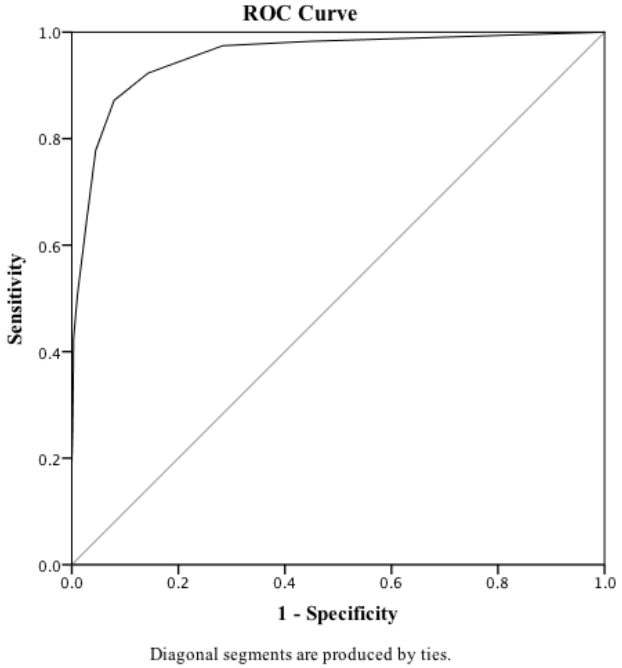



Table 21. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 108 | 83 | 191 |
| **Negative Score (*n*)** | 9 | 497 | 506 |
| **Total (*n*)** | 117 | 580 | 697 |

**Discussion**

As in the previous analyses, no changes were required to the item make-up or the cutoff score for the abbreviated scoring algorithm within the 31-37 month old sample. Sensitivity and specificity estimates were comparable to Study 1A and 1B and the Cervantes et al. (in press) study. Sensitivity ranged from 0.923-0.958; specificity ranged from 0.853-0.857. PPV estimates were slightly higher than estimates found within the younger samples and the Cervantes et al. (in

press) estimate. PPV ranged from 0.526-0.565. NPV estimates continued to remain stable across studies and ranged from 0.982-0.992 in the current study.

Table 22 presents the total scores across individuals scoring true and false positive and true and false negative. As in the 24-30 month study, a total score of 6 was most frequent in the true positive group. Of the true negative group, 63.27% scored a 0 on the total algorithm, and 83.95% scored either a 0 or a 1. In contrast, nearly 80% of the false negative group had a total score of 2 on the abbreviated scoring algorithm. Further, nearly half of the false positive group had a total score of 3 (42.42%); nearly 70% of this group (66.67%) scored a 3 or a 4.

Table 22. Total score on algorithm across participants classified both correctly and incorrectly

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | 16 | 23 | 24 | 39 | 20 | 28 | 15 | 17 | 13 | 4 | 199 |
| **True Negative** | 615 | 201 | 156 | - | - | - | - | - | - | - | - | - | - | 972 |
| **False Positive** | - | - | - | 70 | 40 | 23 | 21 | 7 | 3 | 1 | 0 | 0 | 0 | 165 |
| **False Negative** | 2 | 1 | 10 | - | - | - | - | - | - | - | - | - | - | 13 |

In regards to item endorsement and in line with previous findings, participants who scored false positive were most likely to have ratings of "1" than ratings of "2" across all items. The false positive group had more frequent endorsement of at lease some level of impairment on item [10] Social interactions with others his/her age (76.97% had some level of impairment), item [59] Development of social relationships (64.48% had some level of impairment), and item [4] Engages in repetitive motor movements for no reason (52.12% had some level of impairment). Items that were most frequently endorsed in the true positive group were item [10] Social interactions with others his/her age (93.50% had some level of impairment, 54.48% scored a "2"), item [59] Development of social relationships (88.94% had some level of impairment, 49.75% scored a "2"), and item [8] Maintains eye contact (81.91% had some level of impairment, 36.18% scored a "2"). In the false negative group, ratings of either "1" or "2" were most frequent on item [10] Social interactions with others his/her age (30.77%), item [8]

Maintains eye contact (46.15%), and item [34] Abnormal preoccupation with the parts of object(s) (38.46%). As in previous studies, this trend highlights autism as a primarily social disorder; however, differential diagnosis may be complicated by social impairments present within other related disorders.

Also in line with previous studies, item [10] Social interactions with others his/her age was the most frequently endorsed item within the total sample, and item [34] Abnormal preoccupation with the parts of object(s)was the least frequently endorsed. However, slightly more of the participants with ASD in both the false negative and true positive groups endorsed impairment on item [34] Abnormal preoccupation with the parts of object(s) than in the previous analyses with younger samples. Endorsement of item [4] Engages in repetitive motor movements for no reason remained somewhat stable across age groups within individuals scoring true and false positive and true and false negative on the algorithm.

**Study 1 General Discussion**

Contrary to the hypothesis, no revisions to the algorithm were necessary across the age groups sampled in this study. Psychometric properties were strong across groups of participants who were 17-23 months old, 24-30 months old, and 31-37 months old. In fact, sensitivity and specificity estimates across Studies 1A: 17-23 Months (sensitivity = 0.918-0.963, specificity = 0.851-0.854), 1B: 24-30 Months (sensitivity = 0.950-0.964, specificity = 0.843-0.854), and 1C: 31-37 Months (sensitivity = 0.923-0.958, specificity = 0.853-0.857) were comparable to that of the complete *BISCUIT-Part 1* (sensitivity = 0.94, specificity = 0.87; Konst et al., submitted). Therefore, the abbreviated scoring algorithm reliably differentiates between ASD and no ASD across ages and may be useful in identifying even very young children who require further autism evaluation.

The PPV estimates found within the age studies (i.e., Study 1A: 17-23 Months = 0.442-0.448; Study 1B: 24-30 Months = 0.482-0.531; Study 1C: 31-37 Months = 0.526-0.565) were similar to the PPV of 0.502 found within Cervantes et al. (in press). NPV estimates were comparable as well (i.e., Study 1A: 17-23 Months = 0.989-0.995; Study 1B: 24-30 Months = 0.990-0.994; Study 1C: 31-37 Months = 0.982-0.992; Cervantes et al. [in press] = 0.993). As discussed within Cervantes et al. (in press), it is common to obtain a relative low PPV and a high NPV when studying low prevalence disorders such as ASD. Further, PPV is often low on screening measures because a greater rate of false positives often results from efforts to maximize the number of true positives (Johnson et al., 2007). PPV estimates were higher in increasingly older samples within this study. Several factors may contribute to this. The prevalence of ASD within the 31-37 month sample (15.72%) was higher than in the 24-30 month sample (14.11%) which was higher than in the 17-23 month sample (11.18%). This finding is not surprising given the various symptom onset patterns demonstrated by the ASD popuation; young children who do not exhibit as significant or as many symptoms as needed for clinical diagnosis may meet ASD criteria later in development. Further, symptoms consistent with ASD may become more distinct from symptoms of related but separate DDs with age. This may be particularly true because comparison for item ratings on the *BISCUIT-Part 1* are made against same-aged children. As peers continue to develop more skills across early life, ASD symptoms may become more visible to caregivers.

In terms of item endorsement, those related to social impairments were endorsed at significantly higher rates than RRBIs for children with ASD across the age groups studied. This highlights the idea that ASD at its core is a social disorder. Further, RRBIs have been found less indicative of autism in very early childhood but rather develop across the first several years of

life (Guthrie et al., 2013; Saint-Georges et al., 2010; Wetherby et al., 2004). In line with this finding, slightly more participants with ASD aged 31-37 months old endorsed impairments on item [34] Abnormal preoccupation with the parts of object(s) than in previous studies. Though, endorsement of item [4] Engages in repetitive motor movements for no reason remained stable across age groups. Of note, item [10] Social interactions with others his/her age was the most frequently endorsed item across both individuals with ASD and atypical development in all age groups. This is important; though impairment in peer relationships is integral for an ASD diagnosis, this symptom should not be examined in isolation as many other DDs result in difficulties with peer interactions as well. The same should be said for item [4] Engages in repetitive motor movements for no reason. Item [34] Abnormal preoccupation with the parts of object(s) was the least frequently endorsed item across age groups; however, the item remained a significant predictor of ASD across all regression models. Therefore, although abnormal preoccupation with parts of object(s) may not be a more common symptom within young children with ASD, this symptom may be significantly indicative of autism when present across 17- to 37-month-olds.

# CHAPTER 10. STUDY 2: GENDER

The research literature is inconclusive regarding symptom presentation differences between genders. Although there are discrepancies in the diagnosis of males and females, most current research on younger samples indicates that there are no significant differences in autism symptomology across genders (Frazier et al., 2014; Hiller et al., 2015; Postorino et al., 2015; Rivet & Matson, 2011; Sipes et al., 2011; Van Wijngaarden-Cremers et al., 2014). Therefore, the abbreviated scoring algorithm may perform adequately in detecting autism risk when assessing both male and female participants. To explore the effects of gender, the sample was divided into male and female subsamples. Participants with ASD and participants with atypical development were included in each subsample.

## Study 2A: Male Sample

### Method

**Participants.** The male sample consisted of 5836 participants; 868 were assigned to the ASD group and 4995 to the atypical development group. As in Study 1, the sample was divided in half using random selection. A total of 2927 participants were placed in the exploratory group, and 2936 participants were placed in the revisions/replication group. The exploratory and revisions/replication group did not differ in terms of ethnicity ($X^2$ [3] = 2.71, $p$ = 0.44), age ($F$ [1, 5861] = 0.64, $p$ = 0.42), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 0.17, $p$ = 0.68).

The exploratory group consisted of 439 participants with ASD and 2488 with atypical development. No significant differences were found between participants with ASD and participants with atypical development in age ($F$ [1, 2925] = 3.39, $p$ = 0.07) and ethnicity ($X^2$ [3] = 4.61, $p$ = 0.20). In the revisions/replication group, 2507 participants were assigned to the

atypical development group and 429 participants to the ASD. Significant differences were found between participants with ASD and with atypical development in regards to both ethnicity ($X^2$ [3] = 11.23, $p$ = 0.01) and age ($F$ [1, 2934] = 20.09, $p$ < 0.001, *partial* $\eta^2$ = 0.007). Age differences were small, and ethnicity differences were not expected to influence results for reasons discussed previously. Therefore, analyses were performed with these groups. Demographic information for the exploratory and revisions/replication groups is presented in Table 23.

Table 23. Male Sample Demographics

| | ***Exploratory Group*** | | ***Revisions Group*** | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 2488) | ASD ($n$ = 439) | Atypical Development ($n$ = 2507) | ASD ($n$ = 429) |
| **Age (months)** | | | | |
| *M* (*SD*) | 25.43 (4.65) | 25.88 (4.59) | 25.44 (4.62) | 26.52 (4.46) |
| **Ethnicity** | | | | |
| African American | 35.17% (875) | 39.64% (174) | 35.50% (890) | 42.42% (182) |
| Caucasian | 52.37% (1303) | 48.97% (215) | 53.21% (1334) | 44.52% (191) |
| Hispanic | 3.78% (94) | 2.51% (11) | 3.87% (97) | 4.20% (18) |
| Other/Unspecified | 8.68% (216) | 8.88% (39) | 7.42% (186) | 8.86% (38) |

**Research Design.** The statistics used for Study 1 were repeated with the male sample.

**Results**

**Exploratory Group.** As in Study 1, assumptions were checked prior to running analyses. The logistic regression was then conducted. The model was statistically significant, $\chi^2$ (6) = 1402.85, $p$ < .001. All six items were significant predictors. The model explained 66.7% (Nagelkerke $R^2$) of the variance in the presence of ASD, and the overall correct classification percentage was 92.2% (see Table 24).

Table 24. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 2411 | 77 | 96.9 |
| **ASD** | 150 | 289 | 65.8 |
| **Overall** | - | - | 92.2 |

The subsequent ROC analysis indicated the algorithm had an AUC value of 0.955 (i.e., excellent discriminating ability; Compton et al., 2006). The ROC curve (Figure 8) was used to identify the optimal cutoff score of 3, identical to previous studies. This score yielded a sensitivity of 0.943 and a specificity of 0.845 (see Table 25). PPV was then estimated at 0.518, and NPV was estimated at 0.988 (see Table 26).

Figure 8. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

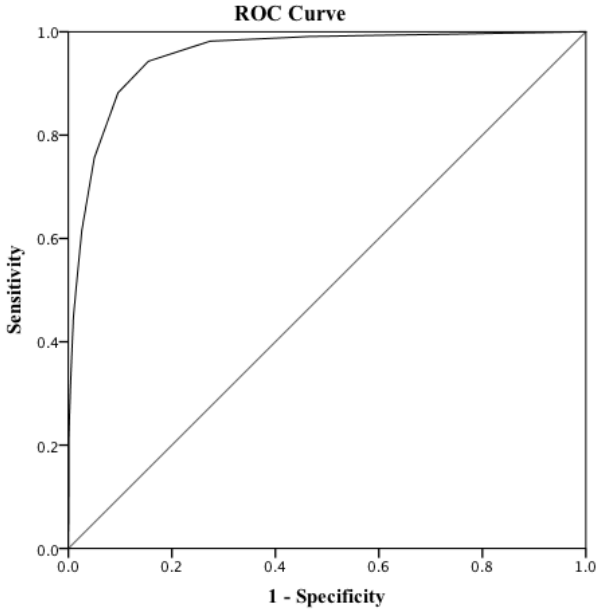



Table 25. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 0.991 | 0.534 |
| 2 | 0.982 | 0.727 |
| **3** | **0.943** | **0.845** |
| 4 | 0.882 | 0.904 |
| 5 | 0.756 | 0.950 |
| 6 | 0.615 | 0.974 |
| 7 | 0.446 | 0.990 |
| 8 | 0.323 | 0.996 |
| 9 | 0.212 | 0.999 |
| 10 | 0.112 | 0.999 |
| 11 | 0.055 | 1.000 |
| 12 | 0.027 | 1.000 |

Table 26. Performance of Algorithm Using a Cutoff Score of 3

| | ASD (*n*) | Atypical Development (*n*) | Total (*n*) |
|---|---|---|---|
| **Positive Score (*n*)** | 414 | 385 | 799 |
| **Negative Score (*n*)** | 25 | 2103 | 2128 |
| **Total (*n*)** | 439 | 2488 | 2927 |

**Revisions/Replication Group.** Again, assumptions were checked and the logistic regression was performed. The model was again statistically significant, $\chi^2$ (6) = 1396.31, $p <$ .001. As in previous analyses, all six items were found to be significant predictors. The model explained 67.0% (Nagelkerke $R^2$) of the variance in presence of ASD and had an overall correct classification percentage of 92.4% (see Table 27).

Table 27. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 2432 | 75 | 97.0 |
| **ASD** | 147 | 282 | 65.7 |
| **Overall** | - | - | 92.4 |

Then, the ROC analysis resulted in an AUC value of 0.957 indicating that the algorithm had excellent discriminating ability within this group (Compton et al., 2006). The ROC curve, shown in Figure 9, was used to assess the cutoff score of 3 found within the exploratory group. Sensitivity was estimated at 0.956, and specificity was 0.836. Further, PPV was 0.499 and NPV was 0.991 in the revisions/replication group (see Table 28).

Figure 9. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

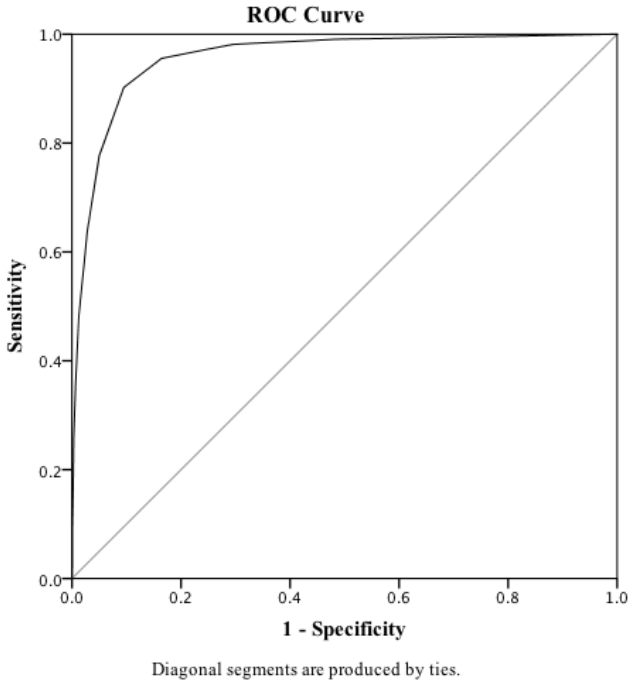



Table 28. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 410 | 412 | 822 |
| **Negative Score (*n*)** | 19 | 2095 | 2114 |
| **Total (*n*)** | 429 | 2507 | 2936 |

**Discussion**

As in previous studies, no changes to the item make-up or cutoff score were necessary to achieve adequate psychometrics (i.e., sensitivity and specificity) for the abbreviated scoring algorithm with the male sample. Sensitivity and specificity estimates ranged from 0.943-0.956 and 0.836-0.845, respectively. PPV ranged from 0.499-0.518. NPV estimates remained stable across studies, ranging from 0.988-0.991 with the male sample. These estimates were similar to those found in Study 1 and Cervantes et al. (in press).

Total scores for the complete male sample across individuals scoring true and false positive and true and false negative are presented in Table 29. Similar to Study 1 results, the

most frequent total score in the true positive group was a 6. This contrasts with results from the participants who scored false positive; nearly 70% of this group had a total score of 3 or 4, and 40.28% scored a 3. Of the true negative group, 62.51% scored a 0 on the total algorithm, and over 85% scored either a 0 or a 1. In contrast, 63.63% of the false negative group had a total score of 2.

Table 29. Total score on algorithm across participants classified both correctly and incorrectly

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | 50 | 109 | 122 | 141 | 109 | 91 | 80 | 61 | 34 | 27 | 824 |
| **True Negative** | 2624 | 952 | 622 | - | - | - | - | - | - | - | - | - | - | 4198 |
| **False Positive** | - | - | - | 321 | 226 | 116 | 79 | 28 | 17 | 3 | 5 | 1 | 1 | 797 |
| **False Negative** | 8 | 8 | 28 | - | - | - | - | - | - | - | - | - | - | 44 |

As in previous studies, children who scored false positive were more likely to have ratings of "1" than "2" when impairment was endorsed. This was true on all items except item [4] Engages in repetitive motor movements for no reason; slightly more of the false positive group had a rating of "2" (27.98%) than "1" (22.84%). Item [10] Social interactions with others his/her age was most commonly endorsed item in the false positive group (71.39% endorsed at least some impairment). In comparison to 31.62% of the false positive group, over half of true positive group were rated a "2" on item [10]. In the true positive group, items [10] Social interactions with others his/her age, [59] Development of social relationships, and [8] Maintains eye contact were the most frequently endorsed social communication and social interaction items. In regards to RRBIs, item [4] Engages in repetitive motor movements for no reason was more commonly endorsed than item [34] Abnormal preoccupation with the parts of object(s). The false negative group was most likely to endorse at least some level of impairment on item [10] Social interactions with others his/her age (38.64% of the group), item [8] Maintains eye contact (34.09%), and item [4] Engages in repetitive motor movements for no reason (22.73%). Over 90% of the true negative group had ratings of "0" on all items except on item [10] Social

interactions with others his/her age (86.23% endorsed no impairment). These findings together suggest that social impairment is imperative in the diagnosis of ASD; however, as previously mentioned, deficient social skills are not exclusive to ASD.

## Study 2B: Female Sample

### Method

**Participants.** The female sample included 2673 participants; 271 of these individuals were assigned to the ASD group and 2402 were assigned to the atypical development group. The sample was again divided in half using random selection to ensure revisions to the algorithm could be made if necessary. A total of 1325 participants were placed in the exploratory group, and 1348 participants were placed in the revisions/replication group. As in previous studies, the exploratory and revisions/replication group did not differ in regards to ethnicity ($X^2$ [3] = 4.05, $p$ = 0.26), age ($F$ [1, 2671] = 1.87, $p$ = 0.17), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 0.05, $p$ = 0.83).

In the exploratory group, there were 136 participants with ASD and 1189 with atypical development. No significant differences were found between participants with ASD and participants with atypical development in age ($F$ [1, 1323] = 3.54, $p$ = 0.06) and ethnicity ($X^2$ [3] = 1.85, $p$ = 0.61). Within the revisions/replication group, 1213 participants had atypical development and 135 participants had ASD. No significant differences were found between participants with ASD and with atypical development in terms of ethnicity ($X^2$ [3] = 4.01, $p$ = 0.26) and age ($F$ [1, 1346] = 2.23, $p$ =0.14). Demographic information for the female sample is presented in Table 30.

Table 30. Female Sample Demographics

| | *Exploratory Group* | | *Revisions Group* | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 1189) | ASD ($n$ = 136) | Atypical Development ($n$ = 1213) | ASD ($n$ = 135) |
| **Age (months)** | | | | |
| *M* (*SD*) | 24.84 (4.60) | 25.63 (4.63) | 25.10 (4.56) | 25.72 (4.59) |
| **Ethnicity** | | | | |
| African American | 40.96% (487) | 44.12% (60) | 36.93% (448) | 43.70% (59) |
| Caucasian | 48.78% (580) | 43.38% (59) | 51.94% (630) | 42.96% (58) |
| Hispanic | 2.94% (35) | 2.94% (4) | 3.46% (42) | 3.71% (5) |
| Other/Unspecified | 7.32% (87) | 9.56% (13) | 7.67% (93) | 9.63% (13) |

**Research Design** The statistics used for Study 1 were repeated with the female sample.

**Results**

**Exploratory Group.** Assumptions of logistic regression were examined prior to running analyses. The logistic regression was then performed and, the regression model was statistically significant, $\chi^2$ (6) = 484.80, $p < .001$. All six items served as significant predictors. The model explained 63.3% (Nagelkerke $R^2$) of the variance in the presence of ASD. The overall correct classification percentage was 93.7% (see Table 31).

Table 31. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 1159 | 30 | 97.5 |
| **ASD** | 54 | 82 | 60.3 |
| **Overall** | - | - | 93.7 |

As in previous studies, the ROC analysis was then conducted and indicated the abbreviated scoring algorithm had an AUC value of 0.961 (i.e., excellent discriminating ability; Compton et al., 2006). The ROC curve (Figure 10) identified an optimal cutoff score of 3, yielding a sensitivity of 0.956 and a specificity of 0.878 (see Table 32). PPV was at 0.473 and NPV was 0.994 (see Table 33).

Figure 10. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

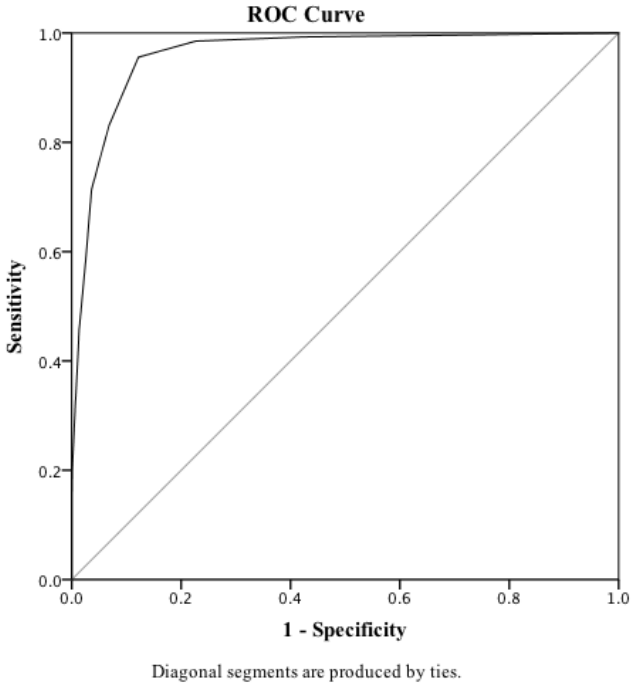



Table 32. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 0.993 | 0.582 |
| 2 | 0.985 | 0.773 |
| **3** | **0.956** | **0.878** |
| 4 | 0.831 | 0.932 |
| 5 | 0.713 | 0.964 |
| 6 | 0.581 | 0.975 |
| 7 | 0.456 | 0.987 |
| 8 | 0.324 | 0.993 |
| 9 | 0.199 | 0.999 |
| 10 | 0.154 | 1.000 |
| 11 | 0.103 | 1.000 |
| 12 | 0.059 | 1.000 |

Table 33. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 130 | 145 | 275 |
| **Negative Score (*n*)** | 6 | 1044 | 1050 |
| **Total (*n*)** | 136 | 1189 | 1325 |

**Revisions/Replication Group.** As in previous analyses, assumptions were first examined. The logistic regression was then conducted. The model was statistically significant, $\chi^2$ (6) = 493.10, $p < .001$; all items of the algorithm were found to be significant predictors. The model explained 64.0% (Nagelkerke $R^2$) of the variance in presence of ASD and had an overall correct classification percentage of 94.5% within the revisions/replication group (see Table 34).

Table 34. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 1194 | 19 | 98.4 |
| **ASD** | 55 | 80 | 59.3 |
| **Overall** | - | - | 94.5 |

The subsequent ROC analysis resulted in an AUC value of 0.956 indicating excellent discriminating ability within the revisions/replication group (Compton et al., 2006). The ROC curve (Figure 11) was then used to assess the cutoff score of 3 found within the exploratory group. Using this cutoff score, sensitivity was 0.933, and specificity was 0.863. PPV was estimated at 0.432 and NPV was at 0.991 in the revisions/replication group (see Table 35).

Figure 11. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

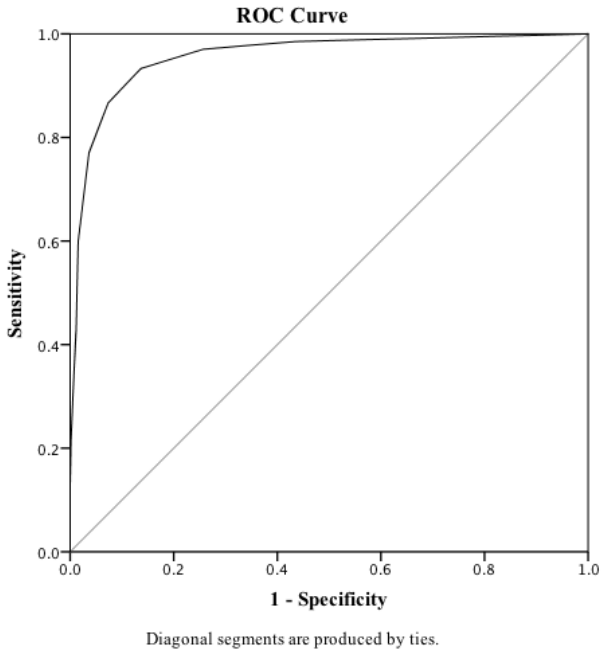


Table 35. Performance of Algorithm Using a Cutoff Score of 3

| | ASD (*n*) | Atypical Development (*n*) | Total (*n*) |
|---|---|---|---|
| **Positive Score (*n*)** | 126 | 166 | 292 |
| **Negative Score (*n*)** | 9 | 1047 | 1056 |
| **Total (*n*)** | 135 | 1213 | 1348 |

**Discussion**

In line with previous studies, no changes to item make-up or cutoff score were necessary to achieve adequate sensitivity and specificity estimates for the female sample. Sensitivity estimates ranged between 0.933-0.956. Specificity ranged from 0.863-0.878. These estimates were comparable to the male sample and to previous studies. PPV ranged from 0.432-0.473, which was slightly lower than the male sample. Similar to the male sample and previous studies, NPV was between 0.991-0.994.

Total scores on the algorithm for the complete female sample across participants scoring true and false positive and true and false negative are presented in Table 36. Compared to a total score of 6 in the male sample, the most frequent total score in the female true positive group was a 5 on the algorithm. In comparison, just over 45% of the false positive group scored a 3, and 72.03% of this group scored either a 3 or 4. Over 65% of the true negative group scored a 0 on the algorithm. Approximately 85% of this group scored either a 0 or a 1. In comparison and similar to previous studies, 60% of the false negative group had a total score of 2.

Table 36. Total score on algorithm across participants classified both correctly and incorrectly

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | 26 | 29 | 41 | 40 | 34 | 30 | 18 | 16 | 10 | 12 | 256 |
| **True Negative** | 1378 | 442 | 271 | - | - | - | - | - | - | - | - | - | - | 2091 |
| **False Positive** | - | - | - | 141 | 83 | 38 | 19 | 15 | 12 | 3 | 0 | 0 | 0 | 311 |
| **False Negative** | 3 | 3 | 9 | - | - | - | - | - | - | - | - | - | - | 15 |

Item endorsement patterns looked similar to the male sample. For the true positive group, item [10] Social interactions with others his/her age was most commonly endorsed. More than half of the true positive group scored a "2" on this item. Items [59] Development of social relationships, [8] Maintains eye contact, and [4] Engages in repetitive motor movements for no reason were also frequently endorsed within this group. Identical to previous studies, participants scoring false positive were more likely to receive ratings of "1" than "2" when impairment was endorsed across all items but item [4] Engages in repetitive motor movements for no reason. The most commonly endorsed items in the false positive group were item [59] Development of social relationships (60.12% of group received a "1", 18.01% received a "2") and item [10] Social interactions with others his/her age (75.88% received a "1"; 36.01% received a '2'). Endorsement trends of the true negative group were also similar to previous studies. Item [10] Social interactions with others his/her age was the most frequently endorsed item (86.66% rated as having no impairment); and, over 90% of the group indicated no impairment across all other items. Within the false negative group, the most commonly endorsed items were item [8] Maintains eye contact (40.00% had at least some level of impairment) and item [59] Development of social relationships (33.33% had at least some impairment). The lowest rated items were items related to RRBIs (i.e., item [4] Engages in repetitive motor movements for no reason and item [34] Abnormal preoccupation with the parts of object[s]). As discussed in previous studies, this endorsement trend emphasizes the concept of ASD as a primarily social disorder particularly in early childhood.

## Study 2 General Discussion

As expected, gender did not significantly impact the performance of the abbreviated scoring algorithm. No changes were necessary. Psychometric properties were strong across

male and female participants. Sensitivity and specificity estimates were 0.943-0.956 and 0.836-0.845 in the male sample and 0.933-0.956 and 0.863-0.878 in the female sample. These results are similar to estimates of the complete *BISCUIT-Part 1* (sensitivity = 0.94, specificity = 0.87; Konst et al., submitted). Therefore, the abbreviated scoring algorithm reliably distinguishes ASD risk across genders.

The PPV estimates found within the male sample were 0.499-0.518 and within the female sample were 0.432-0.473. These estimates were similar to the PPV found in Cervantes et al. (in press) study (PPV=0.502) and were not surprising given the nature of ASD as a low prevalence disorder and the intention of screening instruments (i.e., maximize true positives at the expense of more false positives). PPV estimates were lower in the female sample than the male sample. This is likely due to the lower prevalence of ASD in the female sample (10.14% had ASD) compared to the male sample (14.80% had ASD). The male-to-female ratio within the total sample (approximately 3:1) was slightly lower than the commonly accepted 4:1 ratio within the general population (CDC, 2014). Because these participants were recruited from an early intervention program, this result could be expected. Children enrolled in this program have developmental concerns that are associated with increased risk for ASD (Cervantes et al., in press). Therefore, a majority of females with ASD in this sample likely have comorbidities. As previously mentioned, the gender ratio is less disparate in populations that have ID or behavior problems (Dworzynski et al., 2012). NPV estimates were comparable across studies (i.e., Study 2A: Male = 0.988-0.991; Study 2B: Female = 0.991-0.994) and similar to the original estimate from the Cervantes et al. (in press) study (NPV = 0.993).

Of note, a similar trend of item endorsement across male and female participants was observed. Item [10] Social interactions with others his/her age was the most commonly endorsed

item across the total male and female samples and in participants with ASD scoring true positive (i.e., more than half of the male and female true positive group scored a "2" on this item). Further, endorsement of impairment on items related to social communication and social interaction (i.e., items 10, 59, and 8) occurred more frequently than endorsement on items related to RRBIs (i.e., items 4 and 34) in both the female and male true positive group. The male and female false positive groups were most likely to receive ratings of "1" than "2" on items that were endorsed, and the most commonly endorsed items were related to socialization (i.e., items 10 and 59 in both samples). In the male and female false negative groups, endorsement of impairment was most common on social items as well (i.e., items 10, 8, and 59). Although not a focus of this study, this similar endorsement pattern may support existent research and suggest that symptoms of ASD are similar in males and females at young ages (Postorino et al., 2015; Sipes et al., 2011). However, definite conclusions cannot be drawn as no statistical analyses were performed to address this topic.

## CHAPTER 11. STUDY 3: DEVELOPMENTAL FUNCTIONING

There is significant symptom overlap between ASD and ID, which creates complexity within the assessment process (DiGuiseppi et al., 2010; Peters-Scheffer et al., 2016). A substantial amount of evidence indicates that individuals with ASD and cognitive impairments demonstrate more severe autism symptomology and associated problems (Matson & Shoemaker, 2009; McCarthy et al., 2010; Peters-Scheffer et al., 2016). Sipes and colleagues (2011) showed that this finding was also true for infants and toddlers with ASD and low developmental functioning. Because of the overlap in symptoms between ID and ASD, a higher cutoff score on the abbreviated scoring algorithm may yield improved sensitivity, specificity, and PPV estimates for infants and toddlers with low developmental functioning. In contrast, because of the increased ASD severity in comorbid ASD and ID cases, a lower cutoff score may be necessary to detect autism risk in infants and toddlers with typical developmental functioning.

To identify the algorithm's utility with children of varying developmental levels, the total sample was divided into a low developmental functioning subsample (DQ ≤ 70; i.e., two standard deviations [SDs] below the mean of 100 on the *BDI-2*) and a typical developmental functioning subsample (DQ > 70). DQs from the *BDI-2* were used due to the instability of measures of intellectual functioning (i.e., IQ) in young children; information regarding developmental progression tends to be a more precise assessment in infant and toddlerhood (Matson, Mahan, Hess, & Fodstad, 2010). As in the above studies, each group included participants with ASD and participants with atypical development.

## Study 3A: Low Developmental Functioning Sample

### Method

**Participants.** The low developmental functioning sample consisted of 1704 participants; 601 had ASD and 1103 had atypical development. As in previous analyses, the sample was divided in half. A total of 879 participants were placed in the exploratory group, and 825 participants were placed in the revisions/replication group. The exploratory and revisions/replication group did not differ in terms of gender ($X^2$ [1] = 1.04, $p$ = 0.31), ethnicity ($X^2$ [3] = 6.74, $p$ = 0.08), age ($F$ [1, 1702] = 0.94, $p$ = 0.33), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 0.01, $p$ = 0.92).

The exploratory group included 309 participants with ASD and 570 with atypical development. No significant differences in ethnicity were found between participants with ASD and participants with atypical development ($X^2$ [3] = 1.97, $p$ = 0.58). Significant differences were found in gender ($X^2$ [1] = 4.96, $p$ = 0.03) and age ($F$ [1, 877] = 327.26, $p$ < .001, *partial* $\eta^2$ = 0.015). As in previous studies, gender differences were expected and, age differences were small and not expected to influence results. In the revisions/replication group, 533 participants had atypical development and 292 had ASD. No significant differences were found between participants with ASD and with atypical development in gender ($X^2$ [1] = 2.52, $p$ = 0.11) or ethnicity ($X^2$ [3] = 0.39, $p$ = 0.94). Significant age differences ($F$ [1, 823] = 9.49, $p$ = .002, *partial* $\eta^2$ = 0.011) were found; however, these again were small and not expected to impact results. Demographic information for the exploratory and revisions/replication groups is presented in Table 37.

Table 37. Low Developmental Functioning Sample Demographics

| | *Exploratory Group* | | *Revisions Group* | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 570) | ASD ($n$ = 309) | Atypical Development ($n$ = 533) | ASD ($n$ = 292) |
| **Age (months)** | | | | |
| *M* (*SD*) | 25.09 (4.97) | 26.37 (4.79) | 24.93 (4.87) | 26.00 (4.67) |
| **Gender** | | | | |
| Male | 71.40% (407) | 78.32% (242) | 69.79% (372) | 75.00% (219) |
| Female | 28.60% (163) | 21.68% (67) | 30.21% (161) | 25.00% (73) |
| **Ethnicity** | | | | |
| African American | 43.51% (248) | 39.81% (123) | 42.03% (224) | 43.49% (127) |
| Caucasian | 43.51% (248) | 47.57% (147) | 45.40% (242) | 43.49% (127) |
| Hispanic | 5.26% (30) | 4.21% (13) | 2.63% (14) | 3.09% (9) |
| Other/Unspecified | 7.72% (44) | 8.41% (26) | 9.94% (53) | 9.93% (29) |

**Research Design.** The statistics used for Study 1 were repeated with the low developmental functioning sample.

**Results**

**Exploratory Group.** Assumptions were checked using the same methods described in Study 1. Based on results of the subsequent logistic regression, the model was statistically significant, $\chi^2$ (6) = 553.97, $p < .001$. All six items were significant predictors. The model explained 64.3% (Nagelkerke $R^2$) of the variance in the presence of ASD; the overall correct classification percentage was 84.5% (see Table 38).

Table 38. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 521 | 49 | 91.4 |
| **ASD** | 87 | 222 | 71.8 |
| **Overall** | - | - | 84.5 |

As in previous studies, a ROC analysis was then conducted. The algorithm had an AUC value of 0.925 (i.e., excellent discriminating ability; Compton et al., 2006). The ROC curve (Figure 12) was used to identify the optimal cutoff score of 5. This score yielded a sensitivity of 0.841 and a specificity of 0.870 (see Table 39). Within the exploratory group, PPV was 0.778

and NPV was 0.910 (see Table 40).

Figure 12. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

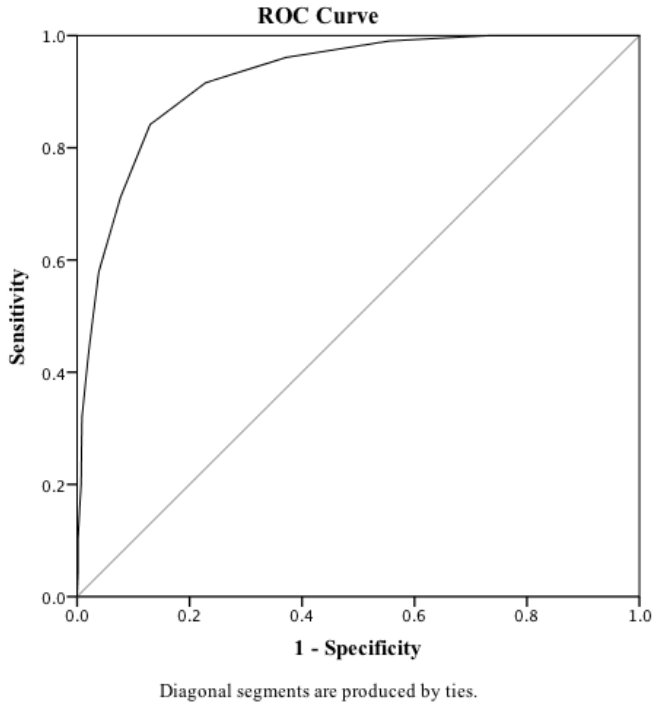



Table 39. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 1.000 | 0.267 |
| 2 | 0.990 | 0.446 |
| 3 | 0.961 | 0.628 |
| 4 | 0.916 | 0.774 |
| **5** | **0.841** | **0.870** |
| 6 | 0.712 | 0.923 |
| 7 | 0.579 | 0.961 |
| 8 | 0.440 | 0.979 |
| 9 | 0.320 | 0.991 |
| 10 | 0.201 | 0.993 |
| 11 | 0.104 | 0.998 |
| 12 | 0.045 | 0.998 |

Table 40. Performance of Algorithm Using a Cutoff Score of 5

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 260 | 74 | 334 |
| **Negative Score (*n*)** | 49 | 496 | 545 |
| **Total (*n*)** | 309 | 570 | 879 |

**Revisions/Replication Group.** Assumptions were checked prior to running analyses. The logistic regression then performed indicated that the model was statistically significant, $\chi^2$ (6) = 544.22, $p < .001$. All items of the algorithm were found to be significant predictors. The model explained 66.4% (Nagelkerke $R^2$) of the variance in the presence of ASD and had an overall correct classification percentage of 84.5% (see Table 41).

Table 41. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 478 | 55 | 89.7 |
| **ASD** | 73 | 219 | 75.0 |
| **Overall** | - | - | 84.5 |

The ROC analysis was then performed and resulted in an AUC value of 0.927 (i.e., excellent discriminating ability; Compton et al., 2006). As in previous studies, the ROC curve (Figure 13) was used to assess the cutoff score of 5 found within the exploratory group. With this cutoff score, sensitivity was estimated at 0.836 and specificity was 0.859. In the revisions/replication group, PPV was estimated at 0.765 and NPV was at 0.905 (see Table 42).

Figure 13. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

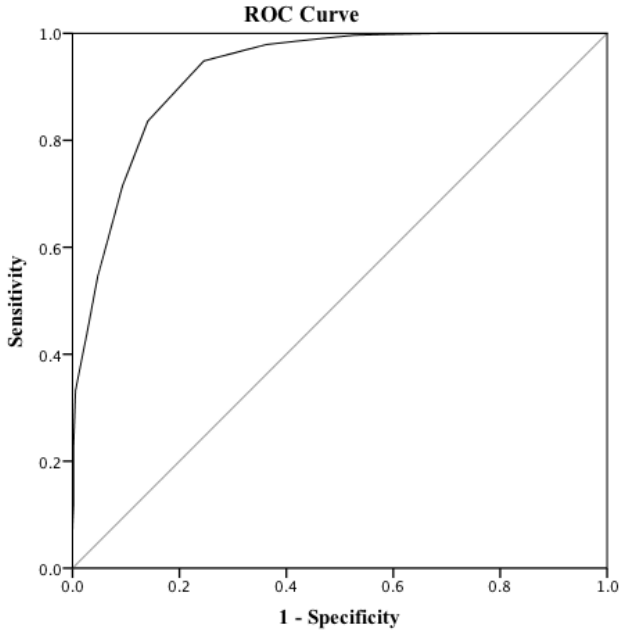


Table 42. Performance of Algorithm Using a Cutoff Score of 5

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 244 | 75 | 319 |
| **Negative Score (*n*)** | 48 | 458 | 506 |
| **Total (*n*)** | 292 | 533 | 825 |

**Discussion**

As hypothesized, a higher cutoff score was necessary to achieve adequate sensitivity and specificity on the algorithm for participants with low developmental functioning. Increasing the cutoff score to 5 resulted in a sensitivity estimate ranging from 0.836-0.841 and a specificity estimate ranging from 0.859-0.870. Though these estimates are lower than in previous studies, examination into an alternative item make-up for the algorithm was not conducted; sensitivity and specificity remained above the recommended 0.80 for screening instruments (Campbell et al., 2016). PPV was higher than in previous studies and ranged from 0.765-0.778. NPV ranged from 0.905-0.910, which is slightly lower than in previous studies.

Total scores on the algorithm across the complete low developmental functioning sample scoring true and false positive and true and false negative are presented in Table 43. Similar to previous studies, the most frequent total score was 6 in the true positive group; however, there were a significant proportion of participants who scored above 6 on the algorithm (67.06% of the true positive group). In comparison, nearly 70% of the false positive group scored a 5 or a 6 on the algorithm. Similar to previous studies, over half of the false negative group had a total score that was one point below the identified cutoff (i.e., a score of 4). Unlike previous studies, only 33.33% of the true negative group scored a 0 on the algorithm. Because the sample consisted of only participants with significant developmental delays, this is not surprising. The overlap between ASD and related but distinct DDs is substantial; this explains not only the need for a

higher cutoff score but also the wide range of total scores across participants who do not have ASD. Of note, if the standard cutoff of 3 identified in all previous studies was utilized, the number of false positives would nearly double. Further, only 79 additional participants (1.12% of the total sample) would be correctly identified as having ASD using the cutoff score of 3.

Table 43. Total score on algorithm across participants classified both correctly and incorrectly

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | - | - | 75 | 91 | 72 | 70 | 68 | 61 | 33 | 34 | 504 |
| **True Negative** | 318 | 188 | 191 | 145 | 112 | - | - | - | - | - | - | - | - | 954 |
| **False Positive** | - | - | - | - | - | 55 | 47 | 20 | 19 | 3 | 3 | 1 | 1 | 149 |
| **False Negative** | 0 | 4 | 14 | 23 | 56 | - | - | - | - | - | - | - | - | 97 |

In regards to item endorsement patterns, some differences were observed in this study compared to previous studies. For example, more frequent ratings of "2" than "1" occurred on all six items within the false positive group. The most commonly endorsed items were item [10] Social interactions with others his/her age (85% of group endorsed as having at least some level of impairment), [8] Maintains eye contact (67.79% had some level of impairment), and [59] Development of social relationships (67.79% had some level of impairment). However, the true positive group demonstrated more frequent ratings of "2" on all items compared to the false positive group. The most frequent items rated as having impairment within this group were item [10] Social interactions with others his/her age (95.63% had some level of impairment, 63.10% had ratings of a "2") and item [59] Development of social relationships (91.47% had some level of impairment, 57.74% had ratings of a "2"). In regards to RRBIs, item [4] Engages in repetitive motor movements for no reason was endorsed more often within the true positive group than item [34] Abnormal preoccupation with the parts of object(s)**.** In previous studies, approximately 90% or more of the true negative groups endorsed no impairments across all items. However, only 70-80% of the true negative group in the low developmental functioning

sample endorsed no impairment. Item [10] Social interactions with others his/her age was most commonly endorsed in this group (29.25% had some level of impairment). Similar to previous studies, items related to socialization were most commonly endorsed in the false negative group. Within this group, endorsements on item [10] Social interactions with others his/her age (65.98% had some level of impairment), item [53] Use of nonverbal communication (58.76% had some level of impairment), and item [59] Development of social relationships (50.52% had some level of impairment) were most frequent.

## Study 3B: Typical Developmental Functioning Sample

### Method

**Participants.** The typical developmental functioning sample included 6832 participants; 538 were assigned to the ASD group and 6294 to the atypical development group. Dividing the sample in half using random selection, a total of 3458 participants were placed in the exploratory group and 3374 participants were placed in the revisions/replication group. The exploratory and revisions/replication group did not differ in terms of gender ($X^2$ [1] = 0.77, $p$ = 0.38), ethnicity ($X^2$ [3] = 2.12, $p$ = 0.55), age ($F$ [1, 6830] = 0.26, $p$ = 0.61), or proportion of participants in ASD and atypical development groups ($X^2$ [1] = 0.11, $p$ = 0.74).

In the exploratory group, there were 276 participants with ASD and 3182 with atypical development. No significant age differences between participants with ASD and participants with atypical development were found ($F$ [1, 3456] = 2.94, $p$ = 0.09); however, differences were found in gender ($X^2$ [1] = 4.92, $p$ = 0.03) and ethnicity ($X^2$ [3] = 7.84, $p$ = 0.05). As discussed previously, neither of these factors was expected to influence results. For the revisions/replication group, 3112 participants had atypical development and 262 had ASD. No significant differences were found between participants with ASD and with atypical

development in ethnicity ($X^2$ [3] = 4.89, $p$ = 0.18). Significant differences were found in gender ($X^2$ [1] = 13.29, $p$ < .001) and age ($F$ [1, 3372] = 5.95, $p$ = 0.02, *partial* $\eta^2$ = 0.002). As previously mentioned, gender differences are not surprising given the nature of ASD and, age differences were small and not expected to effect results. Demographic information is presented in Table 44.

Table 44. Typical Developmental Functioning Sample Demographics

| | ***Exploratory Group*** | | ***Revisions Group*** | |
|---|---|---|---|---|
| | Atypical Development ($n$ = 3182) | ASD ($n$ = 276) | Atypical Development ($n$ = 3112) | ASD ($n$ = 262) |
| **Age (months)** | | | | |
| *M* (*SD*) | 25.31 (4.59) | 25.80 (4.34) | 25.35 (4.55) | 26.07 (4.37) |
| **Gender** | | | | |
| Male | 66.66% (2121) | 73.19% (202) | 67.32% (2095) | 78.24% (205) |
| Female | 33.34% (1061) | 26.81% (74) | 32.68% (1017) | 21.76% (57) |
| **Ethnicity** | | | | |
| African American | 34.73% (1105) | 41.30% (114) | 36.09% (1123) | 42.37% (111) |
| Caucasian | 53.58% (1705) | 46.38% (128) | 53.08% (1652) | 46.18% (121) |
| Hispanic | 3.80% (121) | 2.54% (7) | 3.31% (103) | 3.44% (9) |
| Other/Unspecified | 7.89% (251) | 9.78% (27) | 7.52% (234) | 8.01% (21) |

**Research Design.** The statistics used for Study 1 were repeated with the typical developmental functioning sample.

**Results**

**Exploratory Group.** As in previous studies, assumptions were checked prior to running analyses. The logistic regression was conducted and the model was statistically significant, $\chi^2$ (6) = 1092.53, $p$ < .001. All six items served as significant predictors of the presence of ASD. The model explained 63.5% (Nagelkerke $R^2$) of the variance in the presence of ASD; the overall correct classification percentage was 95.1% (see Table 45).

Table 45. Logistic Regression Classification Table

| Observed | Predicted | | |
|---|---|---|---|
| | No ASD | ASD | Percentage Correct |
| **No ASD** | 3124 | 58 | 98.2 |
| **ASD** | 112 | 164 | 59.4 |
| **Overall** | - | - | 95.1 |

A ROC analysis was then conducted; the algorithm had an AUC value of 0.953 (i.e., excellent discriminating ability within this group; Compton et al., 2006). Next, the ROC curve (Figure 14) was used to identify an optimal cutoff score. In line with previous studies, a cutoff score of 3 was identified and yielded a sensitivity of 0.931 and a specificity of 0.889 (see Table 46). PPV was then estimated at 0.442 and NPV was at 0.993 (see Table 47).

Figure 14. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

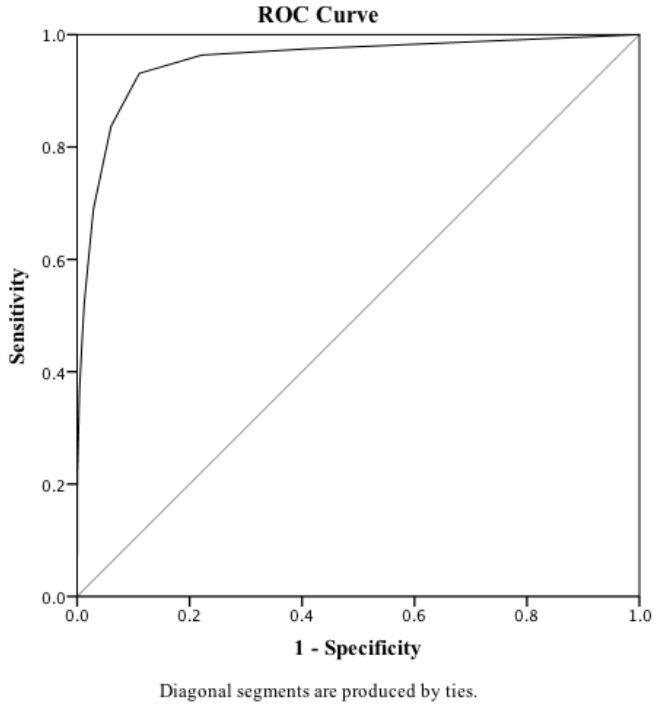


Table 46. Sensitivities and Specificities for Potential Cutoff Scores

| Cutoff Score | Sensitivity | Specificity |
|---|---|---|
| 1 | 0.975 | 0.589 |
| 2 | 0.964 | 0.778 |
| **3** | **0.931** | **0.889** |
| 4 | 0.837 | 0.940 |
| 5 | 0.688 | 0.971 |
| 6 | 0.518 | 0.988 |
| 7 | 0.366 | 0.996 |
| 8 | 0.221 | 0.999 |
| 9 | 0.109 | 1.000 |
| 10 | 0.058 | 1.000 |
| 11 | 0.029 | 1.000 |
| 12 | 0.011 | 1.000 |

Table 47. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 257 | 352 | 609 |
| **Negative Score (*n*)** | 19 | 2830 | 2849 |
| **Total (*n*)** | 276 | 3182 | 3458 |

**Revisions/Replication Group.** Again, assumptions were checked prior to running the replication. The logistic regression was then conducted; the model was statistically significant, $\chi^2$ (6) = 931.39, $p < .001$. All six items were found to be significant predictors. The model explained 57.3% (Nagelkerke $R^2$) of the variance in presence of ASD and had an overall correct classification percentage of 94.6% within the revisions/replication group (see Table 48).

Table 48. Logistic Regression Classification Table

| **Observed** | **Predicted** | | |
|---|---|---|---|
| | **No ASD** | **ASD** | **Percentage Correct** |
| **No ASD** | 3061 | 51 | 98.4 |
| **ASD** | 132 | 130 | 49.6 |
| **Overall** | - | - | 94.6 |

The subsequent ROC analysis resulted in an AUC value of 0.952 (i.e., excellent discriminating ability; Compton et al., 2006). The ROC curve, shown in Figure 15, was then

used to assess the cutoff score of 3 found in the exploratory group. Sensitivity was estimated at 0.916, and specificity was 0.888. PPV was 0.407 and NPV was 0.992 in the revisions/replication group (see Table 49).

Figure 15. ROC Curve Plotting the Sensitivity and 1-Specificity for Each Potential Cutoff Score

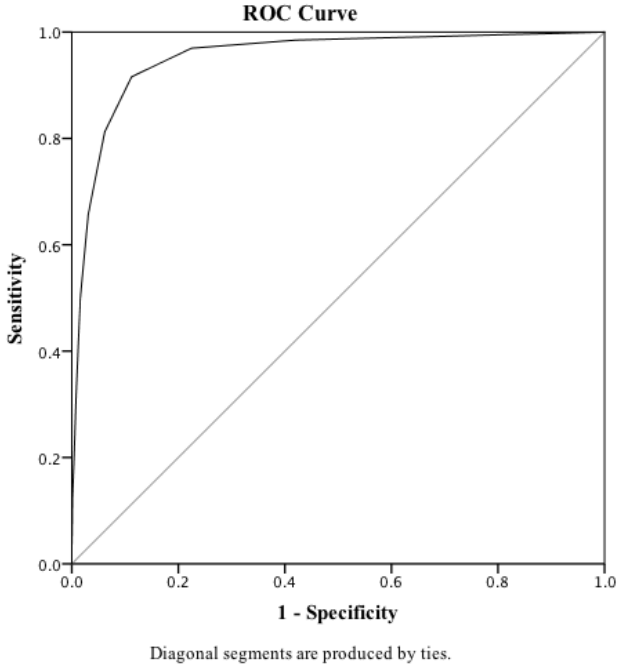



Table 49. Performance of Algorithm Using a Cutoff Score of 3

| | **ASD (*n*)** | **Atypical Development (*n*)** | **Total (*n*)** |
|---|---|---|---|
| **Positive Score (*n*)** | 240 | 350 | 590 |
| **Negative Score (*n*)** | 22 | 2762 | 2784 |
| **Total (*n*)** | 262 | 3112 | 3374 |

**Discussion**

For participants with typical developmental functioning, no changes were necessary to the item make-up or cutoff score of the algorithm to achieve adequate sensitivity and specificity estimates. The sensitivity ranged from 0.916-0.931, and the specificity ranged from 0.888-0.889. This is similar to the estimates found in previous studies and the original Cervantes et al. (in press) study. PPV was found to be 0.442 in the exploratory group and 0.407 in the

revisions/replication group. These estimates are slightly lower than previous studies. NPV ranged from 0.992-0.993, which was fairly commensurate with previous studies.

Table 50 displays the total scores on the algorithm across the complete typical developmental functioning sample (i.e., true and false positive, true and false negative groups). Consistent with previous studies, the most frequent total score was 6 in the true positive group. In the false positive group, 45.16% had a total score of 3; nearly 75% of the group scored either a 3 or a 4 just meeting the cutoff for a positive score. Likewise and as in previous studies, over half of the false negative group had a total score of 2. Over 65% of the true negative group had a total score of 0, and 87.45% of the group scored either a 0 or a 1 on the algorithm.

Table 50. Distribution of total scores on algorithm across participants classified both correctly and incorrectly

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **True Positive** | - | - | - | 53 | 82 | 88 | 90 | 71 | 51 | 30 | 16 | 11 | 5 | 497 |
| **True Negative** | 3684 | 1206 | 702 | - | - | - | - | - | - | - | - | - | - | 5592 |
| **False Positive** | - | - | - | 317 | 197 | 99 | 51 | 23 | 10 | 3 | 2 | 0 | 0 | 702 |
| **False Negative** | 11 | 7 | 23 | - | - | - | - | - | - | - | - | - | - | 41 |

In regards to item endorsement patterns, the trend within the typical developmental functioning sample appeared similar to that found in Study 1 and Study 2. When impairment was endorsed on the items of the algorithm, the false positive group obtained more ratings of "1" than "2"; this was true for every item but item [4] Engages in repetitive motor movements for no reason. The most commonly endorsed items were item [10] Social interactions with others his/her age (74.07% had some level of impairment) and item [59] Development of social relationships (57.41% had some level of impairment). These items were also the most frequently endorsed items for the true positive group; though, they were endorsed at much higher rates. In the true positive group, 92.76% had some level of impairment rated on item [10] Social interactions with others his/her age (48.49% were rated a "2"); 87.53% had some level of

impairment on item [59] Development of social relationships (38.43% were rated a "2"). Similar to previous studies, all socialization items (i.e., items 8, 10, 53, and 59) were endorsed at higher rates than items related to RRBIs (i.e., items 4 and 34). In the false negative group, socialization items were also more frequently endorsed. This was particularly true for item [10] Social interactions with others his/her age (39.02% had some level of impairment) and item [8] Maintains eye contact (36.59% had some level of impairment). Identical to previous studies, more than 90% of true negative group were rated as having no impairment on all items except item [10] Social interactions with others his/her age (87.59% had a rating of "0").

**Study 3 General Discussion**

As hypothesized, participant developmental level affected the performance of the abbreviated scoring algorithm. The algorithm had strong sensitivity and specificity estimates with use of the original cutoff score of 3 on the typical developmental functioning sample. Within this group, sensitivity ranged from 0.916-0.931 and specificity ranged from 0.888-0.889. Like in previous studies, these estimates are similar to the complete *BISCUIT-Part 1* estimates (sensitivity = 0.94, specificity = 0.87; Konst et al., submitted).

Changes to the cutoff score but not to the item make-up were necessary to achieve adequate sensitivity and specificity for use with children with low developmental functioning. Using a cutoff score of 5, sensitivity ranged from 0.836-0.841 and specificity ranged from 0.859-0.870 within this sample. While the specificity estimates were similar to the complete *BISCUIT-Part 1*, sensitivity was lower. The increased cutoff score and lower sensitivity estimate within this group are not surprising. Due to the substantial overlap in symptoms, it is more difficult to discern ASD from DDs associated with low developmental functioning (such as ID). Children who experience significant cognitive deficits are also likely to demonstrate significant social and

communication deficits. Because ASD symptoms must be in excess of the symptoms accounted for by developmental impairment to meet criteria for diagnosis (DiGuiseppi et al., 2010; Peters-Scheffer et al., 2016), the need for a higher cutoff score within this sample makes conceptual sense.

The PPV estimates found across developmental functioning samples were also disparate. PPV within the low developmental functioning sample ranged from 0.765-0.778, and PPV within the typical developmental functioning sample ranged from 0.407-0.442. As in previous studies, the sample with the higher prevalence rate of ASD (i.e., the low developmental functioning sample, 35.27% had ASD) had higher PPV estimates. Within the typical developmental functioning sample, 7.87% had ASD. The difference in ASD prevalence rates across samples may be explained by the relationship between ID/GDD and ASD; more significant impairments are associated with an increase in the prevalence of ASD (DiGuiseppi et al., 2010; Turygin, Matson, & Adams, 2014). Although children in the low developmental functioning group did not necessarily have an ID diagnosis at the time of assessment, these participants will likely meet criteria for ID later in life given their substantial developmental delays (Tirosh & Jaffe, 2011). NPV estimates were slightly lower within the low developmental functioning sample (i.e., Study 3A: Low Developmental Functioning = 0.905-0.910; Study 3B: Typical Developmental Functioning = 0.992-0.993) but still strong. Differences in NPV are likely due to ASD prevalence differences between samples as well.

In regards to item endorsement, higher total scores on the algorithm were observed across groups in the low developmental functioning sample compared to the typical developmental functioning sample. In addition, there were more frequent endorsements of impairment (i.e., ratings of "1" or "2") across groups in the low developmental functioning sample on all items.

Again, these results are to be expected given the association between ASD and more significant developmental/cognitive impairment (DiGuiseppi et al., 2010; Turygin et al., 2014). The most commonly endorsed items within both low and typical developmental functioning samples were related to socialization rather than RRBIs. Within the true positive groups specifically, the most common items where impairment was endorsed were the same across samples (i.e., items [10] Social interactions with others his/her age and [59] Development of social relationships).

# CHAPTER 12. CONCLUSION

The abbreviated scoring algorithm in its original form performed well across age groups and genders, and with participants with typical developmental functioning. Sensitivity estimates within these subgroups (i.e., 17-23 months, 24-30 months, 31-37 months, females, males, and participants with typical developmental functioning) were all above 0.90; specificity was between 0.836-0.889. PPV across all groups were similar to that of other available ASD screening instruments, and NPV was above 0.90 in all groups. The ability of the abbreviated scoring algorithm to hold up across these groupings evidences the clinical utility of the algorithm to serve as a screener.

The abbreviated scoring algorithm did not perform adequately in its original form with participants with low developmental functioning. However, when the cutoff score was increased by two points, sensitivity and specificity rose to above 0.80. Using the higher cutoff score, PPV was highest in this group (0.765-0.778). This was likely due to the higher prevalence of ASD in the low developmental functioning sample as PPV and NPV are heavily influenced by disorder prevalence (Barton et al., 2012; Johnson et al., 2007). NPV estimates were above 0.90. However, in regards to feasibility of utilizing an alternate cutoff score based upon developmental functioning, the author understands that it is not always possible to estimate a child's developmental level prior to ASD screening. When an estimate of developmental level is available, the cutoff score of 5 should be used with children with substantial deficits. In cases where developmental level is unknown, clinicians are encouraged to be mindful of the limitations of the abbreviated scoring algorithm for use with children with substantial delays. Though, a referral for a more comprehensive developmental evaluation would only benefit a child demonstrating significant deficits consistent with having a DQ equal to or more than 2 SDs

from the mean. Following a more extensive diagnostic workup, a more specific diagnosis may be able to be provided to these children who would likely score false positive on the algorithm in its original form. Thereafter, more specific services and supports, more information regarding prognosis, and perhaps more detailed genetic counseling could be available for the child and his or her family (Moeschler, Shevell, & Committee on Genetics, 2014).

Administering the abbreviated scoring algorithm within the two-pronged assessment process discussed within the Purpose section of this study may also add psychometric strength to the screening process. To review, this would involve first administering the algorithm to all participants. Those participants who screen positive on the algorithm would be asked to complete the full 62-item measure. Then, the participants who score at-risk on both the algorithm and the complete *BISCUIT-Part 1* would be referred to a specialty diagnostic clinic. This process would enable those who are not at-risk for ASD to avoid the more extensive 62-item screen by scoring negative on the algorithm. Further, administering the complete *BISCUIT-Part 1* following a positive score on the algorithm would result in a reduction of false positives.

Interestingly, no changes were needed to the item make-up of the abbreviated scoring algorithm across groups. As discussed in the Cervantes et al. (in press) study, items of the algorithm correspond well with the *DSM-5* ASD criteria and the research regarding early ASD symptom emergence which explains the items' flexibility across age groups, genders, and developmental levels. For example, social and communication symptoms have been found to emerge earlier than RRBIs in young children who are later diagnosed with ASD (Guthrie et al., 2013; Saint-Georges et al., 2010; Wetherby et al., 2004); and, four of the six items in the algorithm represent the social domain whereas only two items are related to RRBIs. Further, symptoms consistent with these social items on the algorithm (e.g., joint attention deficits, eye

contact deficits, limited/lack of interest in interaction and sharing interests) have been consistently found to best discern ASD from other DDs across the early detection research (Cervantes et al., in press; Saint-Georges et al., 2010; Wetherby et al., 2004; Zwaigenbaum et al., 2007).

Various trends in item endorsement in this study were also noteworthy. First, clinicians should be aware that as a child's total score approaches the cutoff, there is a greater possibility that the child will be categorized as a false positive (i.e., scoring just above the cutoff) or false negative (i.e., scoring just below the cutoff). Further, children who score false positive on the screener were found to receive more ratings of "1" than "2" in most groupings (except in the low developmental functioning sample). In regards to individual item patterns, item [10] Social interactions with others his/her age was highly endorsed across both children with ASD and with atypical development. In fact, within the revisions/replication group of the 17-23 month sample, item [10] did not serve as a significant predictor in the regression model. This is important for differential diagnosis. Based on the results of these studies, difficulties with peer interactions do not appear exclusive to ASD. Though, severe impairments in peer interactions (i.e., ratings of a "2" on the item) seemed more indicative of ASD. On the contrary, item [34] Abnormal preoccupation with the parts of object(s) was uncommonly endorsed throughout the studies but was consistently a significant predictor in the regression models. Therefore, while the social items and item [4] Engages in repetitive motor movements for no reason were endorsed at higher rates, abnormal preoccupation with parts of objects may be significantly suggestive of autism when present.

Although results of the current study demonstrate the strength of the abbreviated scoring algorithm to serve as an ASD screener, further research is necessary to support its use. First, the

study should be replicated using a typically developing sample to extend findings to populations that would be screened within primary care settings. There is a significant need for a time efficient, psychometrically sound measure for PCP use as formal ASD screening often fails to be conducted by PCPs and the clinical judgment of PCPs regarding ASD risk has been found less accurate than formal screening tools (Crais et al., 2014; Johnson et al., 2007). Therefore, this research examining the algorithm's utility within alternate settings is imperative. Further, the use of the abbreviated scoring algorithm as a screener should be studied prospectively. Because the current study was retrospective, all items of the *BISCUIT-Part 1* were given to caregivers. Administering all items may prime caregivers to respond to the items within the algorithm a certain way. This may also be the case when administering the *BISCUIT-Part 1* within an assessment battery (e.g., conducting a developmental evaluation with the autism screener). As such, further research examining how the algorithm performs on its own is necessary. Within these prospective studies, particular emphasis should also be placed on the systematic study of parent and clinician acceptability of the screening tool (i.e., the abbreviated scoring algorithm) and the screening procedure (e.g., if the two-pronged assessment process was used, the efficient referral to specialty clinics). This evaluation would help determine if the introduction of the algorithm as a screener would serve its purpose in increasing feasibility while reducing the burden of ASD screening on professionals (Cervantes et al., in press).

Based upon these results, the abbreviated scoring algorithm appears to show promise in serving as a measure for early detection of autism risk across ASD subgroups. The algorithm demonstrated psychometric strength across various subgroups comparable to other available ASD screeners (Zwaigenbaum et al., 2015). This is particularly true given the large and representative sample used in this study; the proportion of children who did and did not have

ASD was reflective of the high-risk population examined, adding strength to these results. To review, the algorithm also offers many unique benefits in addition to psychometric strength (Cervantes et al., in press). First, the items of the *BISCUIT* are rated by caregivers on a 3-point Likert scale. Parent-report measures offer ease in administration and the ability to assess symptoms across a range of settings and time (Barton et al., 2012). Further, the 3-point Likert scale may offer greater response flexibility compared to the yes/no designation of other autism screeners (e.g., *M-CHAT*; Crais et al., 2014). This flexibility may be integral when differentiating variation in typical development and/or symptoms of a related but distinct DD from autism. In regards to concerns about potential caregiver comprehension difficulties (Crais et al., 2014), the *BISCUIT*'s appendix is available and was developed to improve understanding. Lastly, as discussed, the algorithm may offer the time efficiency needed for effective practice by early intervention providers and PCPs (Cervantes et al., in press).

With further research, the algorithm could be a strong and efficient ASD screening tool for early childhood professionals across settings (e.g., primary care, early intervention). Continued advancements to the abbreviated scoring algorithm, other ASD screening instruments, and to screening procedures would be beneficial on a multitude of levels. First, ASD screening likely leads to earlier diagnosis. Achieving diagnostic clarity earlier in the child's development lends more time for caregiver adjustment, stress management, and psychoeducation. Acceptance of an ASD diagnosis and knowledge regarding ASD would lead to more efficient and appropriate service identification and coordination. When children are enrolled in evidence-based treatments (such as EIBI) at earlier ages, prognosis improves. Researchers have shown that children enrolled in EIBI at younger ages make larger gains in adaptive functioning, IQ, and ASD symptoms (Granpeesheh et al., 2009; Smith, Klorman, & Mruzek, 2015). Therefore,

earlier intervention can promote an improved quality of life for the child and his or her family. Further, the potential for significant long-term financial savings for the family and the government after early evidence-based intervention has been cited (Chasson, Harris, & Neely, 2007). Because early diagnosis has such a significant impact on the future of a child and his or her family, efforts towards improving early detection of ASD need to continue.

## REFERENCES


Al-Qabandi, M., Gorter, J. W., & Rosenbaum, P. (2011). Early autism detection: Are we ready for routine screening? *Pediatrics*, *128*(1), e211–e217.

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders (3rd ed.)*. Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders (3rd ed. - Revised)*. Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, (4th ed.)*. Washington, DC: Author.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders-Text Revision, (4th ed., text rev.)*. Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, (5th ed.)*. Washington, DC: Author.

Arunyanart, W., Fenick, A., Ukritchon, S., Imjaijitt, W., Northrup, V., & Weitzman, C. (2012). Developmental and autism screening: A survey across six states. *Infants & Young Children*, *25*(3), 175–187.

Asperger, H., & Frith, U. (1991). "Autistic psychopathy" in childhood. In *Autism and Asperger syndrome* (pp. 37–92). New York, NY, US: Cambridge University Press.

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., & Drew, A. (2000). A screening instrument for autism at 18 Months of Age: A 6-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(6), 694–702.

Baird, G., Charman, T., Cox, A., Baron-Cohen, S., Swettenham, J., Wheelwright, S., & Drew, A. (2001). Screening and surveillance for autism and pervasive developmental disorders. *Archives of Disease in Childhood*, *84*(6), 468–475.

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry*, *161*(6), 839–843.

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

Barton, M. L., Dumont-Mathieu, T., & Fein, D. (2012). Screening young children for autism spectrum disorders in primary practice. *Journal of Autism and Developmental Disorders*, *42*(6), 1165–1174.

Begeer, S., Mandell, D., Wijnker-Holmes, B., Venderbosch, S., Rem, D., Stekelenburg, F., & Koot, H. (2012). Sex differences in the timing of identification among children and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(5), 1151–1156.

Beighley, J. S., Matson, J. L., Rieske, R. D., Jang, J., Cervantes, P. E., & Goldin, R. L. (2013). Comparing challenging behavior in children diagnosed with autism spectrum disorders according to the DSM-IV-TR and the proposed DSM-5. *Developmental Neurorehabilitation*, *16*(6), 375–381.

Beighley, J. S., Matson, J. L., Rieske, R. D., Konst, M. J., & Tureck, K. (2014). Differences in communication skills in toddlers diagnosed with Autism Spectrum Disorder according to the DSM-IV-TR and the DSM-5. *Research in Autism Spectrum Disorders*, *8*(2), 74–81.

Bender, L. (1959). Autism in children with mental deficiency. *American Journal of Mental Deficiency*, *64*(1), 81–86.

Bettelheim, B. (1967). *The Empty Fortress*. New York, NY: The Free Press.

Bleuler, E. (1913). Autistic thinking. *American Journal of Insanity*, *1913*(69), 873–886.

Campbell, J. M., Scheil, K. A., & Hammond, R. K. (2016). Screening methods. In J. L. Matson (Ed.), *Handbook of Assessment and Diagnosis of Autism Spectrum Disorder* (pp. 65–86). Springer International Publishing.

Centers for Disease Control and Prevention (2007). *Prevalence of Autism Spectrum Disorders (ASDs).* Retrieved February 23, 2016, from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5601a2.htm

Centers for Disease Control and Prevention (2014). *Prevalence of Autism Spectrum Disorders (ASDs).* Retrieved February 23, 2016, from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6302a1.htm

Cervantes, P. E., Matson, J. L., & Goldin, R. L. (2016). Diagnosing ASD in very early childhood. In J. L. Matson (Ed.), *Handbook of Assessment and Diagnosis of Autism Spectrum Disorder* (pp. 157–173). Springer International Publishing.

Cervantes, P. E., Matson, J. L., & Peters, W. J. (2016). Abbreviated scoring algorithm for the Baby and Infant Screen for Children with aUtIsm Traits. *Developmental Neurorehabilitation*. http://doi.org/10.1080/17518423.2016.1211189

Chasson, G., Harris, G., & Neely, E. (2007). Cost comparison of early intensive behavioral intervention and special education for children with autism. *Journal of Family Sciences, 16,* 401-413.

Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007). Autism spectrum disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry*, *48*(2), 128–138.

Cheslack-Postava, K., & Jordan-Young, R. M. (2012). Autism spectrum disorders: Toward a gendered embodiment model. *Social Science & Medicine*, *74*(11), 1667–1674.

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the Modified Checklist for Autism in low risk toddlers. *Pediatrics*, *131*(4), e1121–e1127.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, *98*(2), 394–409.

Crais, E. R., McComish, C. S., Humphreys, B. P., Watson, L. R., Baranek, G. T., Reznick, J. S., … Earls, M. (2014). Pediatric healthcare professionals' views on autism spectrum disorder screening at 12–18 months. *Journal of Autism and Developmental Disorders*, *44*(9), 2311–2328.

Daniels, A. M., & Mandell, D. S. (2013). Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism*. http://doi.org/10.1177/1362361313480277

Darrou, C., Pry, R., Pernon, E., Michelon, C., Aussilloux, C., & Baghdadli, A. (2010). Outcome of young children with autism: Does the amount of intervention influence developmental trajectories? *Autism*, *14*(6), 663-677.

De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child & Adolescent Psychiatry*, *7*(3), 131–136.

DiGuiseppi, C., Hepburn, S., Davis, J. M., Fidler, D. J., Hartway, S., Lee, N. R., … Robinson, C. (2010). Screening for autism spectrum disorders in children with Down syndrome. *Journal of Developmental and Behavioral Pediatrics*, *31*(3), 181–191.

Dosreis, S., Weiner, C. L., Johnson, L., & Newschaffer, C. J. (2006). Autism spectrum disorder screening and management practices among general pediatric providers. *Journal of Developmental and Behavioral Pediatrics*, *27*(2), S88–94.

Dumont-Mathieu, T., & Fein, D. (2005). Screening for autism in young children: The Modified Checklist for Autism in Toddlers (M-CHAT) and other measures. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(3), 253–262.

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., … Shih, A. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. *Autism Research*, *8*(5), 473-476.

Dworzynski, K., Ronald, A., Bolton, P., & Happé, F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(8), 788–797.

Eaves, L. C., Wingert, H., & Helena, H. H. (2006). Screening for autism: Agreement with diagnosis. *Autism*, *10*(3), 229–242.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology*, *38*(3), 439–450.

Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., … Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders: Global epidemiology of autism. *Autism Research*, *5*(3), 160–179.

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Dawson, G., … Volkmar, F. R. (2000). Practice parameter: Screening and diagnosis of autism Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology*, *55*(4), 468–479.

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders*, *33*(4), 365–382.

Fombonne, E., Quirke, S., & Hagen, A. (2009). Prevalence and interpretation of recent trends in rates of pervasive developmental disorders. *McGill Journal Of Medicine*, *12*(2), 99–107.

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and cognitive characteristics of females and males with autism in the simons simplex collection. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(3), 329–340.e3.

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., Constantino, J., … Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(1), 28–40.e3.

Gabriels, R. L., Cuccaro, M. L., Hill, D. E., Ivers, B. J., & Goldson, E. (2005). Repetitive behaviors in autism: relationships with associated clinical features. *Research in Developmental Disabilities*, *26*(2), 169–181.

Gibbs, V., Aldridge, F., Chandler, F., Witzlsperger, E., & Smith, K. (2012). Brief report: An exploratory study comparing diagnostic outcomes for autism spectrum disorders under DSM-IV-TR with the proposed DSM-5 revision. *Journal of Autism and Developmental Disorders*, *42*(8), 1750–1756.

Gillis, J. M. (2009). Screening Practices of Family Physicians and Pediatricians in 2 Southern States. *Infants and Young Children*, *22*(4), 321–331.

Goin-Kochel, R. P., Mackintosh, V. H., & Myers, B. J. (2006). How many doctors does it take to make an autism spectrum diagnosis? *Autism*, *10*(5), 439–451.

Granpeesheh, D., Dixon, D. R., Tarbox, J., Kaplan, A. M., & Wilke, A. E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *3*(4), 1014–1022.

Gray, K. M., & Tonge, B. J. (2005). Screening for autism in infants and preschool children with developmental delay. *Australian and New Zealand Journal of Psychiatry*, *39*(5), 378–386.

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation: Stability of early diagnosis and symptoms in ASD. *Journal of Child Psychology and Psychiatry*, *54*(5), 582–590.

Hiller, R. M., Young, R. L., & Weber, N. (2015). Sex differences in pre-diagnosis concerns for children later diagnosed with autism spectrum disorder. *Autism*. http://doi.org/10.1177/1362361314568899

Horovitz, M., & Matson, J. L. (2013). The Baby and Infant Screen for Children with aUtism traits-Part 2: The development of age-based cutoffs. *Developmental Neurorehabilitation*, *18*(6), 349-356.

Hyman, S. L., & Johnson, J. K. (2012). Autism and pediatric practice: Toward a medical home. *Journal of Autism and Developmental Disorders*, *42*(6), 1156–1164.

Inglese, M. D., & Elder, J. H. (2009). Caring for children with autism spectrum disorder, part I: Prevalence, etiology, and core features. *Journal of Pediatric Nursing*, *24*(1), 41–48.

Johnson, C. P., Myers, S. M., & and the Council on Children With Disabilities. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1183–1215.

Kalb, L. G., Law, J. K., Landa, R., & Law, P. A. (2010). Onset patterns prior to 36 Months in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(11), 1389–1402.

Kamio, Y., Haraguchi, H., Miyake, A., & Hiraiwa, M. (2015). Brief report: Large individual variation in outcomes of autistic children receiving low-intensity behavioral interventions in community settings. *Child and Adolescent Psychiatry and Mental Health*, *9*(1), 6.

Kanner, L. (1943). Autistic disturbances of affective contact. *Acta Paedopsychiatrica*, *35*(4), 100–136.

Kanner, L. (1944). Early infantile autism. *The Journal of Pediatrics*, *25*(3), 211–217.

Kirkovski, M., Enticott, P. G., Hughes, M. E., Rossell, S. L., & Fitzgerald, P. B. (2016). Atypical neural activity in males but not females with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 954–963.

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., … Fein, D. (2008). The Modified Checklist for Autism in Toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(5), 827–839.

Konst, M. J., Matson, J. L., & Matheis, M. (submitted). The Baby and Infant Screen for Children with aUtism Traits- Part 1 (BISCUIT-Part 1): A DSM-5 update.

Kreiser, N. L., & White, S. W. (2014). ASD in females: Are we overstating the gender difference in diagnosis? *Clinical Child and Family Psychology Review*, *17*(1), 67–84.

Lai, M.-C., Lombardo, M. V., Pasco, G., Ruigrok, A. N. V., Wheelwright, S. J., Sadek, S. A., … MRC AIMS Consortium. (2011). A behavioral comparison of male and female adults with high functioning autism spectrum conditions. *PLoS ONE*, *6*(6), e20835.

Lampi, K. M., Lehtonen, L., Tran, P. L., Suominen, A., Lehti, V., Banerjee, P. N., … Sourander, A. (2012). Risk of autism spectrum disorders in low birth weight and small for gestational age infants. *The Journal of Pediatrics*, *161*(5), 830–836.

Landa, R., & Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry*, *47*(6), 629–638.

Leonard, H., Dixon, G., Whitehouse, A. J. O., Bourke, J., Aiberti, K., Nassar, N., … Glasson, E. J. (2010). Unpacking the complex nature of the autism epidemic. *Research in Autism Spectrum Disorders*, *4*(4), 548–554.

Levy, Y. (2011). Developmental delay revisited. *Developmental Disabilities Research Reviews*, *17*(2), 180–184.

Limperopoulos, C., Bassan, H., Sullivan, N. R., Soul, J. S., Robertson, R. L., Moore, M., … du Plessis, A. J. (2008). Positive screening for autism in ex-preterm infants: Prevalence and risk factors. *Pediatrics*, *121*(4), 758–765.

Lockyer, L., & Rutter, M. (1970). A five- to fifteen-year follow-up study of infantile psychosis: IV. Patterns of cognitive ability. *British Journal of Social and Clinical Psychology*, *9*(2), 152–163.

Lord, C., & Bishop, S. L. (2015). Recent advances in autism research as reflected in DSM-5 criteria for autism spectrum disorder. *Annual Review of Clinical Psychology*, *11*(1), 53-70.

Lord, C., & Luyster, R. (2006). Early diagnosis of children with autism spectrum disorders. *Clinical Neuroscience Research*, *6*(3-4), 189–194.

Lotter, V. (1966). Epidemiology of autistic conditions in young children. *Social Psychiatry*, *1*(3), 124–135.

LoVullo, S. V., & Matson, J. L. (2012). Development of a critical item algorithm for the Baby and Infant Screen for Children with aUtIsm Traits. *Research in Autism Spectrum Disorders*, *6*(1), 378–384.

Mandell, D., & Lecavalier, L. (2014). Should we believe the Centers for Disease Control and Prevention's autism spectrum disorder prevalence estimates? *Autism*, *18*(5), 482–484.

Mandell, D., & Mandy, W. (2015). Should all young children be screened for autism spectrum disorder? *Autism*, *19*(8), 895–896.

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S., … Kirby, R. S. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health*, *99*(3), 493–498.

Matheis, M., & Matson, J. L. (2015). Autism spectrum disorder screening refusal rates: Findings from a statewide early intervention program. *Journal of Developmental and Physical Disabilities*, *27*(6), 755-770.

Matson, J. L., Boisjoli, J. A., Hess, J. A., & Wilkins, J. (2010). Factor structure and diagnostic fidelity of the Baby and Infant Screen for Children with aUtIsm Traits–Part 1 (BISCUIT–part 1). *Developmental Neurorehabilitation*, *13*(2), 72–79.

Matson, J. L., Boisjoli, J., & Wilkins, J. (2007). *The Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT).* Baton Rouge, LA: Disability Consultants, LLC.

Matson, J. L., Hattier, M. A., & Williams, L. W. (2012). How does relaxing the algorithm for autism affect DSM-V prevalence rates? *Journal of Autism and Developmental Disorders*, *42*(8), 1549–1556.

Matson, J. L., & Kozlowski, A. M. (2011). The increasing prevalence of autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(1), 418–425.

Matson, J. L., Kozlowski, A. M., Hattier, M. A., Horovitz, M., & Sipes, M. (2012). DSM-IV vs DSM-5 diagnostic criteria for toddlers with autism. *Developmental Neurorehabilitation*, *15*(3), 185–190.

Matson, J. L., Mahan, S., Hess, J. A., & Fodstad, J. C. (2010). Effect of developmental quotient on symptoms of inattention and impulsivity among toddlers with Autism Spectrum Disorders. *Research in Developmental Disabilities*, *31*(2), 464–469.

Matson, J. L., & Nebel-Schwalm, M. S. (2007). Comorbid psychopathology with autism spectrum disorder in children: An overview. *Research in Developmental Disabilities*, *28*(4), 341–352.

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities*, *30*(6), 1107–1114.

Matson, J. L., Wilkins, J., & Fodstad, J. C. (2011). The Validity of the Baby and Infant Screen for Children with aUtIsm Traits: Part 1 (BISCUIT: Part 1). *Journal of Autism and Developmental Disorders*, *41*(9), 1139–1146.

Matson, J. L., Wilkins, J., & Gonzalez, M. (2008). Early identification and diagnosis in autism spectrum disorders in young children and infants: How early is too early? *Research in Autism Spectrum Disorders*, *2*(1), 75–84.

Matson, J. L., Wilkins, J., Sevin, J. A., Knight, C., Boisjoli, J. A., & Sharp, B. (2009). Reliability and item content of the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT): Parts 1–3. *Research in Autism Spectrum Disorders*, *3*(2), 336–344.

Matson, J. L., Wilkins, J., Sharp, B., Knight, C., Sevin, J. A., & Boisjoli, J. A. (2009). Sensitivity and specificity of the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT): Validity and cutoff scores for autism and PDD-NOS in toddlers. *Research in Autism Spectrum Disorders*, *3*(4), 924–930.

Matson, J. L., Worley, J. A., Fodstad, J. C., Chung, K.-M., Suh, D., Jhin, H. K., … Furniss, F. (2011). A multinational study examining the crosscultural differences in reported symptoms of autism spectrum disorders: Israel, South Korea, the United Kingdom, and the United States of America. *Research in Autism Spectrum Disorders*, *5*(4), 1598–1604.

Mattila, M.-L., Kielinen, M., Linna, S.-L., Jussila, K., Ebeling, H., Bloigu, R., … Moilanen, I. (2011). Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: An epidemiological study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(6), 583–592.

McCarthy, J., Hemmings, C., Kravariti, E., Dworzynski, K., Holt, G., Bouras, N., & Tsakanikos, E. (2010). Challenging behavior and co-morbid psychopathology in adults with intellectual disability and autism spectrum disorders. *Research in Developmental Disabilities*, *31*(2), 362–366.

McDonald, L. (2006). Investigation of global developmental delay. *Archives of Disease in Childhood*, *91*(8), 701–705.

McGovern, C. W., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry*, *46*(4), 401–408.

McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 368–383.

Moeschler, J. B., Shevell, M., & Committee on Genetics. (2014). Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics*, *134*(3), e903–e918.

Newborg, J. (2005). *Battelle Developmental Inventory (2nd ed.).* Itasca, NY: Riverside

Newschaffer, C. J., Croen, L. A., Daniels, J., Giarelli, E., Grether, J. K., Levy, S. E., … Windham, G. C. (2007). The epidemiology of autism spectrum disorders. *Annual Review of Public Health*, *28*(1), 235–258.

Oosterling, I. J., Wensing, M., Swinkels, S. H., van der Gaag, R. J., Visser, J. C., Woudenberg, T., … Buitelaar, J. K. (2010). Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach. *Journal of Child Psychology and Psychiatry*, *51*(3), 250–258.

Ozonoff, S., Heung, K., Byrd, R., Hansen, R., & Hertz-Picciotto, I. (2008). The onset of autism: patterns of symptom emergence in the first years of life. *Autism Research*, *1*(6), 320–328.

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, *96*(1), 3–14.

Peters-Scheffer, N., Didden, R., & Lang, R. (2016). Intellectual disability. In J. L. Matson (Ed.), *Comorbid Conditions Among Children with Autism Spectrum Disorders* (pp. 283–300). Springer International Publishing.

Phadke, V. K., Bednarczyk, R. A., Salmon, D. A., & Omer, S. B. (2016). Association between vaccine refusal and vaccine-preventable diseases in the United States: A review of measles and pertussis. *JAMA*, *315*(11), 1149–1158.

Pizur-Barnekow, K., Muusz, M., McKenna, C., O'Connor, E., & Cutler, A. (2012). Service coordinators' perceptions of autism-specific screening and referral practices in early intervention. *Topics in Early Childhood Special Education*. http://doi.org/10.1177/0271121412463086

Postorino, V., Fatta, L. M., De Peppo, L., Giovagnoli, G., Armando, M., Vicari, S., & Mazzone, L. (2015). Longitudinal comparison between male and female preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(7), 2046-2055.

Rivet, T. T., & Matson, J. L. (2011). Review of gender differences in core symptomatology in autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(3), 957–976.

Robins, D. L., Adamson, L. B., Barton, M., Jr, J. E. C., Dumont-Mathieu, T., Dworkin, P. H., … Vivanti, G. (2016). Universal autism screening for toddlers: recommendations at odds. *Journal of Autism and Developmental Disorders*, 1–3.

Robins, D. L., Casagrande, K., Barton, M., Chen, C.-M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, *133*(1), 37–45.

Robins, D. L., & Dumont-Mathieu, T. M. (2006). Early screening for autism spectrum disorders: update on the modified checklist for autism in toddlers and other measures. *Journal of Developmental & Behavioral Pediatrics*, *27*(2), S111–S119.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *31*(2), 131–144.

Rogers, S. J. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research*, *2*(3), 125–137.

Russell, G., Steer, C., & Golding, J. (2011). Social and demographic factors that influence the diagnosis of autistic spectrum disorders. *Social Psychiatry and Psychiatric Epidemiology*, *46*(12), 1283–1293.

Rutter, D. M., & Bartak, L. (1971). Causes of infantile autism: Some considerations from recent research. *Journal of Autism and Childhood Schizophrenia*, *1*(1), 20–32.

Rutter, M. (1968). Concepts of autism: A review of research. *Journal of Child Psychology and Psychiatry*, *9*(1), 1–25.

Rutter, M. (1978). Diagnosis and definition of childhood autism. *Journal of Autism and Childhood Schizophrenia*, *8*(2), 139–161.

Rutter, M., & Schopler, E. (1987). Autism and pervasive developmental disorders: Concepts and diagnostic issues. *Journal of Autism and Developmental Disorders*, *17*(2), 159–186.

Saemundsen, E., Ludvigsson, P., Hilmarsdottir, I., & Rafnsson, V. (2007). Autism spectrum disorders in children with seizures in the first year of life: A population-based study. *Epilepsia*, *48*(9), 1724–1730.

Saint-Georges, C., Cassel, R. S., Cohen, D., Chetouani, M., Laznik, M.-C., Maestro, S., & Muratori, F. (2010). What studies of family home movies can teach us about autistic infants: A literature review. *Research in Autism Spectrum Disorders*, *4*(3), 355–366.

Schaafsma, S. M., & Pfaff, D. W. (2014). Etiologies underlying sex differences in autism spectrum disorders. *Frontiers in Neuroendocrinology*, *35*(3), 255–271.

Schendel, D., & Bhasin, T. K. (2008). Birth weight and gestational age characteristics of children with autism including a comparison with other developmental disabilities. *Pediatrics*, *121*(6), 1155–1164.

Shah, K. (2001). What Do Medical Students Know about Autism? *Autism*, *5*(2), 127–133.

Siegel, B. (2004). *The Pervasive Developmental Disorders Screening Test-II (PDDST-II).* San Antonio, TX: Psychological Corporation.

Sipes, M., Matson, J. L., Worley, J. A., & Kozlowski, A. M. (2011). Gender differences in symptoms of autism spectrum disorders in toddlers. *Research in Autism Spectrum Disorders*, *5*(4), 1465–1470.

Siu, A. L., and the US Preventive Services Task Force (USPSTF), Bibbins-Domingo, K., Grossman, D. C., Baumann, L. C., Davidson, K. W., … Pignone, M. P. (2016). Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement. *JAMA*, *315*(7), 691.

Smith, T., Klorman, R., & Mruzek, D. W. (2015). Predicting outcome of community-based early intensive behavioral intervention for children with autism. *Journal of Abnormal Child Psychology*, *43*(7), 1271–1282.

Smith, I. C., Reichow, B., & Volkmar, F. R. (2015). The Effects of DSM-5 Criteria on Number of Individuals Diagnosed with Autism Spectrum Disorder: A Systematic Review. *Journal of Autism and Developmental Disorders*, *45*(8), 2541–2552.

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Brief report: Screening Tool for Autism in Two-year-olds (STAT): Development and preliminary data. *Journal of Autism and Developmental Disorders*, *30*(6), 607–612.

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of Autism and Developmental Disorders*, *34*(6), 691–701.

Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the Screening Tool for Autism in Two-Year-Olds (STAT) for children under 24 months: An exploratory study. *Autism*, *12*(5), 557–573.

Tager-Flusberg, H. (2010). The origins of social impairments in autism spectrum disorder: Studies of infants at risk. *Neural Networks*, *23*(8-9), 1072–1076.

Taylor, B. (2006). Vaccines and the changing epidemiology of autism. *Child: Care, Health and Development*, *32*(5), 511–519.

Taylor, L. E., Swerdfeger, A. L., & Eslick, G. D. (2014). Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies. *Vaccine*, *32*(29), 3623–3629.

Tirosh, E., & Jaffe, M. (2011). A pediatric perspective: Global developmental delay and mental retardation. *Developmental Disabilities Research Reviews*, *17*(2), 85–92.

Turner, L. M. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism*, *10*(3), 243–265.

Turner, M. (1999). Annotation: Repetitive behaviour in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry*, *40*(6), 839–849.

Turygin, N., Matson, J. L., & Adams, H. (2014). Prevalence of co-occurring disorders in a sample of adults with mild and moderate intellectual disabilities who reside in a residential treatment setting. *Research in Developmental Disabilities*, *35*(7), 1802–1808.

Turygin, N., Matson, J. L., Beighley, J., & Adams, H. (2013). The effect of DSM-5 criteria on the developmental quotient in toddlers diagnosed with autism spectrum disorder. *Developmental Neurorehabilitation*, *16*(1), 38–43.

Van Wijngaarden-Cremers, P. J. M., van Eeten, E., Groen, W. B., Van Deurzen, P. A., Oosterling, I. J., & Van der Gaag, R. J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: A systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, *44*(3), 627–635.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review*, *30*(4), 387–399.

Waterhouse, L., & Gillberg, C. (2014). Why autism must be taken apart. *Journal of Autism and Developmental Disorders*, *44*(7), 1788–1792.

Werner, E., Dawson, G., Munson, J., & Osterling, J. (2005). Variation in early developmental course in autism and its relation with behavioral outcome at 3–4 years of age. *Journal of Autism and Developmental Disorders*, *35*(3), 337–350.

Wetherby, A. M., Watt, N., Morgan, L., & Shumway, S. (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *Journal of Autism and Developmental Disorders*, *37*(5), 960–975.

Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders*, *34*(5), 473–493.

Wiggins, L. D., Baio, J. O. N., & Rice, C. (2006). Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. *Journal of Developmental & Behavioral Pediatrics*, *27*(2), S79–S87.

Williams, J. (2006). Screening for autism spectrum disorders: What is the evidence? *Autism*, *10*(1), 11–35.

Wing, L., Gould, J., & Gillberg, C. (2011). Autism spectrum disorders in the DSM-V: Better or worse than the DSM-IV? *Research in Developmental Disabilities*, *32*(2), 768–773.

Worley, J. A., & Matson, J. L. (2012). Comparing symptoms of autism spectrum disorders using the current DSM-IV-TR diagnostic criteria and the proposed DSM-V diagnostic criteria. *Research in Autism Spectrum Disorders*, *6*(2), 965–970.

Zander, E., & Bölte, S. (2015). The new DSM-5 impairment criterion: A challenge to early autism spectrum disorder diagnosis? *Journal of Autism and Developmental Disorders*, *45*(11), 3634–3643.

Zwaigenbaum, L. (2010). Advances in the early detection of autism. *Current Opinion in Neurology*, *23*(2), 97–102.

Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., ... Kasari, C. (2015). Early screening of autism spectrum disorder: recommendations for practice and research. *Pediatrics*, *136*, S41-59.

Zwaigenbaum, L., & Stone, W. (2006). Early screening for autism spectrum disorders in clinical practice settings. In T. Charman & W. Stone (Eds.), *Social and Communication Development in Autism Spectrum Disorders: Early Identification, Diagnosis, and Intervention* (pp. 88–113). Guilford Press.

Zwaigenbaum, L., Thurm, A., Stone, W., Baranek, G., Bryson, S., Iverson, J., … Sigman, M. (2007). Studying the emergence of autism spectrum disorders in high-risk infants: Methodological and practical issues. *Journal of Autism and Developmental Disorders*, *37*(3), 466–480.

# APPENDICES

Bobby Jindal
GOVERNOR

Kathy Kliebert
SECRETARY

## State of Louisiana

Department of Health and Hospitals

December 4, 2013

Dr. Johnny L. Matson
Department of Psychology
Louisiana State University
324 Audubon Hall
Baton Rouge, LA 70803

Via email: johnmatson@aol.com

Re: Autism in Early Childhood

Dear Dr. Matson:

Thank you for submitting the above-referenced proposal. We have taken into advisement information provided in the proposal package. We find that all areas of concerns were clarified and the project has been approved by Expedited Review.

The IRB approves the project for the purposes of investigating developmental patterns and differences in atypically developing children with and without autism spectrum disorders. If you should desire to conduct additional research using the data collected under this project, that proposal much be submitted separately to the IRB for review.

I am requesting that the Principal investigator report to the DHH IRB any emergent problems, serious adverse reactions, or changes to protocol that may affect the status of the investigation and that no such changes be instituted prior to DHH IRB review, except where necessary in order to eliminate immediate hazards. The investigator also agrees to periodic review of this project by the DHH IRB at intervals appropriate to the degree of risk to assure that the project is being conducted in compliance with the DHH IRB's understanding and recommendations.

If I can be of any further assistance to you, please feel free to contact me.

Sincerely,

Nell W. Allbritton, MPA
Director, Institutional Review Board
Department of Health and Hospitals
628 North 4th Street, Third Floor
Baton Rouge, Louisiana 70802
(225) 342-4169
nell.allbritton@la.gov

Bienville Building • 628 N. 4th Street • P.O. Box 629 • Baton Rouge, Louisiana 70821-0629
Phone #: 225/342-9509 • Fax #: 225/342-5568 • WWW.DHH.LA.GOV
"An Equal Opportunity Employer"

# Application for Exemption from Institutional Oversight

Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/irb

Unless qualified as meeting the specific criteria for exemption from Institutional Review Board (IRB) oversight, ALL LSU research/ projects using living humans as subjects, or samples, or data obtained from humans, directly or indirectly, with or without their consent, must be approved or exempted in advance by the LSU IRB. This Form helps the PI determine if a project may be exempted, and is used to request an exemption.

– Applicant,Please fill out the application in its entirety and include the completed application as well as parts A-F, listed below, when submitting to the IRB. Once the application is completed, please submit two copies of the completed application to the IRB Office or to a member of the Human Subjects Screening Committee. Members of this committee can be found at http://research.lsu.edu/CompliancePoliciesProcedures/InstitutionalReviewBoard%28IRB%29/item24737.html

– A Complete Application Includes All of the Following:
(A) Two copies of this completed form and two copies of parts B thru F.
(B) A brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts 1&2)
(C) Copies of all instruments to be used.
*If this proposal is part of a grant proposal, include a copy of the proposal and all recruitment material.
(D) The consent form that you will use in the study (see part 3 for more information.)
(E) Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project, including students who are involved with testing or handling data, unless already on file with the IRB. Training link: (http://phrp.nihtraining.com/users/login.php)
(F) IRB Security of Data Agreement: (http://research.lsu.edu/files/item26774.pdf)

**1) Principal Investigator:** Dr. Johnny L. Matson **Rank:** Clinical Director

**Dept:** Clinical Psychology **Ph:** (225) 578-8745 **E-mail:** johnmatson@aol.com

**2) Co Investigator(s):** please include department, rank, phone and e-mail for each
*If student, please identify and name supervising professor in this space

IRB# E8292 LSU Proposal #
✓ Complete Application
✓ Human Subjects Training

**3) Project Title:** Autism in Early Childhood

Study Exempted By:
Dr. Robert C. Mathews, Chairman
Institutional Review Board
Louisiana State University
203 B-1 David Boyd Hall
225-578-8692 | www.lsu.edu/irb
Exemption Expires: 4/30/2016

**4) Proposal? (yes or no)** no **If Yes, LSU Proposal Number**

Also, if YES, either ◯ **This application completely matches the scope of work in the grant**
**OR** ◯ **More IRB Applications will be filed later**

**5) Subject pool** (e.g. Psychology students) Young children assessed for developmental delay via EarlySteps

*Circle any **"vulnerable populations" to be used**: (children <18; the mentally impaired, pregnant women, the ages, other). Projects with incarcerated persons cannot be exempted.

**6) PI Signature** J Matson **Date** 4-19-13 (no per signatures)

** **I certify my responses are accurate and complete.** If the project scope or design is later changes, I will resubmit for review. I will obtain written approval from the Authorized Representative of all non-LSU institutions in which the study is conducted. I also understand that it is my responsibility to maintain copies of all consent forms at LSU for three years after completion of the study. If I leave LSU before that time the consent forms should be preserved in the Departmental Office.

**Screening Committee Action:** Exempted ✓ Not Exempted ____ Category/Paragraph 4

**Signed Consent Waived?** Yes / No

**Reviewer** Mathews **Signature** Robert C Mathews **Date** 5/1/13

# VITA

Paige Cervantes is originally from Levittown, New York. She received her bachelor's degree in psychology with an emphasis in applied behavior analysis from Binghamton University in 2012. She then enrolled in Louisiana State University's clinical psychology doctoral program under the supervision of Dr. Johnny L. Matson. Her clinical and research interests include the assessment and treatment of autism spectrum disorder and related developmental disabilities. She is now completing her doctoral internship at Kennedy Krieger Institute, Johns Hopkins School of Medicine. Paige anticipates graduating with her Ph.D. in August 2017. She will continue her work in the autism field thereafter as a postdoctoral fellow in the Autism Spectrum Disorder Clinical and Research Program at the NYU Langone Medical Center's Child Study Center.