Improving Early Identification of Young Toddlers with Autism Spectrum

Disorders (ASD) Using the Autism Detection in Early Childhood (ADEC)

Yong-Hwee Nah, B.SocSci (Hons), M.A (Applied Psych)

Flinders University

Faculty of Social and Behavioural Sciences

School of Psychology

A thesis submitted to Flinders University for the degree of Doctor of Philosophy

February, 2015

TABLE OF CONTENTS

SUMMARY

Screening for Autism Spectrum Disorders (ASD) is a crucial first step to improve early identification of children who might be considered at risk of the disorder and in need of further assessment, intervention and services. Given early identification and intervention of ASD can dramatically improve outcome for people with ASD (Dawson & Burner, 2011), there is a pressing need to identify children with ASD as early as practical (Reichow, 2012). However, the clinical and etiologic heterogeneity of young children with ASD pose a challenge for clinicians and paediatricians to identify these children in their practices and thus these professionals require appropriate tools and training if they are going to be able to identify these children successfully.

In this thesis, I presented three studies which investigated the psychometric properties of an observation screening measure, the Autism Detection in Early Childhood (ADEC; Young, 2007) in the early identification of young children with possible ASD. At the commencement of this research the ADEC was relatively new, and despite promising data, it has not been subjected to scientific rigour. Study 1 provided a comprehensive psychometric validation of the ADEC as a screening tool for ASD. This study compared 70 children with Autistic Disorder with 57 children with other developmental disorders and 64 typically developing children on the ADEC. The data showed that the ADEC is an effective screening tool that can be used to identify children with ASD ranging from 12 to 36 months.

Study 2 compared the predictive validity data of the ADEC against a well-established screening tool, the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1998), in relation to diagnostic classifications, symptom severity and functioning level at 2 and 6 years following initial assessment.

Participants were 55 children aged 19–42 months at initial assessment who were followed up 2 and 6 years after their initial assessment. Results indicated that both tools performed similarly when predicting long term outcomes such as diagnostic status and overall adaptive functioning level. Although these findings need to be replicated with additional and larger samples, this study extends our understanding of the psychometric properties of both the ADEC and the CARS.

In study 3, I (a) examined the frequency and pattern of diagnostic features detected using the ADEC in children aged from 12 to 71 months, and (b) identified the critical items at each age stage. This provided the basis for the development of a brief version of the ADEC (BADEC) that is valid for different age groups. The dataset used 251 participants with a DSM-5 diagnosis of ASD and 206 non-ASD. Analyses supported the use of those critical items (e.g., response to name and gaze switch) identified across most of the age groups to form one BADEC version for all age groups, albeit with different cutoff scores. The brief version for the different age groups had acceptable internal consistency, correlated with the full version, and mostly had sensitivity and specificity exceeding 80%. The BADEC versions' total scores (with the exception of the 60-71 months group) predicted DSM-5 ASD classification just as well as the more time-intensive ADOS and ADI-R diagnostic tools. However, these results would need to be replicated with larger samples.

The studies in this thesis represent the first step in understanding the psychometric properties and usefulness of using the ADEC, in the early detection of young children with possible ASD. The data from this thesis support the use of the ADEC to be a quick and suitable screening tool by clinicians and pediatricians to help them to identify these children in their practice settings.

DECLARATION

I certify that this thesis does not contain any material which has been accepted for the award of any other degree or diploma; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text of the thesis or notes.

Yong-Hwee Nah, B.SocSci (Hons), M.A (Applied Psych)

27th February, 2015

ACKNOWLEDGEMENTS

First and foremost, thank you to my supervisors, Robyn and Neil. I have the privilege and benefit of working with two of the best minds that I could possible wish for. I have indeed learnt much as a practitioner and a researcher from both of you throughout this process.

Special mention goes to my fellow researchers in ASD who have kept me sane and alive throughout this journey: Darren – I'm proud to call you my first Aussie *mate*; Clare and Anna – my 'room' mates who have provided insightful conversations and laughter in our office; and Bec who reminded me how baking can be a great stress-reliever.

Lastly, I thank God for my most lovely wife, Juliet. She has been my amazing help-mate and my joy. Thank you for tolerating my 'ASD' traits and giving up your SG family, friends and food just to be with me here in Adelaide. You are the best wife that I can ever ask for.

To God by the glory.

STATEMENT OF CO-AUTHORSHIP

Chapter 2.

Nah, Y. H., Young, R., Brewer, N., & Berlingeri, G. (2014). Autism Detection in

Early Childhood (ADEC): Reliability and validity data for a Level 2

screening tool for Autistic Disorder. *Psychological Assessment, 26,* 215-

226. doi:10.1037/a0034472

Nah, Y. H. conceptualized and designed the study, participated in data collection,

analyzed the data, and drafted the initial manuscript; Young, R. and Brewer, N.

conceptualized the study and critically reviewed and edited the manuscript;

Berlingeri, G. conceptualized the study and participated in data collection; and all

authors approved the final manuscript as submitted.

Chapter 3.

Nah, Y. H., Young, R. L., & Brewer, N. (2014). Using the Autism Detection in

Early Childhood (ADEC) and Childhood Autism Rating Scales (CARS) to

predict long term outcomes in children with Autism Spectrum Disorders.

Journal of Autism and Developmental Disorders, 44, 2301-2310.

doi:10.1007/s10803-014-2102-1

Nah, Y. H. conceptualized and designed the study, participated in data collection,

analyzed the data, and drafted the initial manuscript; Young, R. and Brewer, N.

conceptualized the study and critically reviewed and edited the manuscript; and all

authors approved the final manuscript as submitted.

CHAPTER 1

General Introduction

Autism Spectrum Disorder (ASD) is characterised by impairments in

two domains: (1) social communication, and (2) presence of restricted and

repetitive behaviours and interests (*DSM-5;* American Psychiatric Association,

2013). Recent prevalence studies have suggested that ASD may be present in as

many as 11 out of every 1,000 children (Centers for Disease Control and

Prevention, 2012; Elsabbagh et al., 2012; Fombonne, 2009), making ASD one

of the most frequently occurring childhood neurodevelopmental disorders.

Clinicians and paediatricians are likely to see an increase in children presenting

with the possibility of ASD and will need appropriate tools and training to

identify them.

Screening for ASD is a crucial first step for improving early

identification of children who might be considered at risk of the disorder and in

need of further assessment, intervention and services. After screening, the

second step of a comprehensive evaluation may be carried out in order to

accurately rule in or out an ASD or other developmental problem. Given early

identification and intervention of ASD can dramatically improve outcome of

ASD (Dawson & Burner, 2011), there is a pressing need to identify children

with ASD as early as possible (Reichow, 2012).

Previous research has consistently documented a gap between the age at

which children with ASD *can* be identified and the age at which they *are*

identified. For instance, parents of children with ASD may typically report

abnormalities in the development of their child during the first 2 years when

responding to a widely established parental interview tool, the Autism

Diagnostic Interview-Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003).

Moreover, they may even report specific early autistic symptoms (Chawarska,

Klin, Paul, & Volkmar, 2007), though children often do not get diagnosed until

when they are older. For example, Sivberg (2003) reported a delay of 20 to 60

months between parental suspicion and diagnosis by a medical professional,

depending on the severity of the disorder and autism classification. Young,

Brewer, and Pattison (2003) also reported similar delays. In another study by

Shattuck et al. (2009), children could be identified with an ASD as late as 6

years old (median 5.7 years).

Why are children often only diagnosed long after their initial signs

present? There are many reasons for this delay, but of interest here is the delay

caused by clinicians' lack of familiarity with the critical early signs or *red flags*

of ASD (Crais et al., 2014). While these *red flags* are often integrated in the

currently available screening tools, research suggests the tools currently

available are inadequate, time consuming or costly, thereby minimising the

uptake of ASD screening (AAP, 2003; Sices, Feudtner, McLaughlin, Drotar, &

Williams, 2003; Honigfeld & McKay 2006; Gura, Champagne, & Blood-

Siegfried, 2011). Therefore, there is a critical need for further research to

develop and identify valid screening tools that can identify the presence of an

ASD in young children.

The goal of this thesis is to examine the psychometric properties of an

ASD-specific screening tool, the Autism Detection in Early Childhood (ADEC;

Young, 2007). In this introductory chapter, the definition of ASD is provided,

followed by consideration of the screening and diagnosis of ASD and a review

of existing screening and diagnostic measures. Lastly, I describe the ADEC in

greater detail and contrast it with other available screening tools. The following three chapters provide evidence that the ADEC can be a quick and suitable screening tool to help clinicians and paediatricians to identify young children presenting with possible ASD in their practice settings through the three studies conducted. Study 1 provides reliability and validity data for the ADEC. Study 2 presents data on the efficacy of the ADEC and another well-established screening tool, the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1998), for predicting long term outcomes in children with ASD. Study 3 examines how the behaviours operationalised in the ADEC items change as the children with and without ASD get older and identifies a brief version that is suitable for identifying children with possible ASD across the age range from 12 to 71 months.

The Definition of ASD: DSM-IV-TR versus DSM-5

Under the previous *Diagnostic and Statistical Manual of Mental Disorders: 4th Edition Text Revision* (*DSM-IV-TR;* American Psychiatric Association, 2000), Pervasive Developmental Disorder (also commonly referred to as ASD) is a term which refers to a range of lifelong neurodevelopmental conditions comprising Autistic Disorder (AD), pervasive developmental disorder - not otherwise specified (PDD-NOS), and Asperger syndrome (AS) characterised by impairments in three domains: (1) reciprocal social interaction, (2) communication, and (3) presence of restricted and repetitive behaviours/interests.

ASD can range from a severe form, AD (or classic autism), to Asperger's syndrome; often considered to be the less severe variant of the disorder. If a child has many of the symptoms of either of these disorders, but

does not meet the requisite number of criteria for either, a diagnosis of pervasive developmental disorder not otherwise specified (PDD-NOS) may be considered.

AD, PDD-NOS, and AS are the three most commonly occurring ASDs (Fombonne, 2009). During the early 1990s AD was thought to occur in only 0.5 of every 1000 children, with estimates increasing to about 1.2 of every 1000 children during the early 2000s (Fombonne, 2005). Yet recent studies have suggested that the prevalence rate of AD alone is estimated to be as many as two to four out of every 1,000 children (Fombonne, 2009; Parner et al., 2011), with prevalence estimates rising steeply in the last two decades (Fombonne, 2005).

DSM-IV-TR. Best-estimate clinical (BEC) diagnoses of specific ASD (AD, PDD-NOS, and AS) based on the DSM-IV-TR have been used as the diagnostic gold standard, especially for diagnosing children under the age of 5 years (Volkmar, Chawarska, & Klin, 2005). Clinicians derive the BEC DSM-IV-TR diagnoses by using all available information and assessment results to generate independent diagnoses of autism, PDD-NOS and non-autism spectrum disorders. BEC diagnoses using DSM-IV-TR or International Classification of Diseases (10th rev.; ICD-10; World Health Organization, 1993) criteria are commonly used by autism researchers (e.g., Chawarska et al., 2007; Lord et al., 2006) to categorise participants by diagnosis and have been shown to be reliable (Klin, Lang, Cicchetti, & Volkmar, 2000) and generally stable over time, even for children under 3 years of age at their initial diagnosis (Moore & Goodson, 2003; Stone et al., 1999).

However, the applicability of the DSM-IV-TR approaches to the diagnosis of autism in infants and very young children has been questioned on several grounds: for instance, some of the criteria (e.g., impaired conversational

ability, stereotyped and repetitive use of language, and inflexible adherence to routines and rituals; Charman & Baird, 2002; Cox et al., 1999; Eaves & Ho, 2004; Lord, 1995; Moore & Goodson, 2003; Stone et al., 1999) are not clearly applicable to infants. It is likely that these behaviours are more of a reflection of the developmental level of toddlers, with many of them likely to emerge later in the developmental course of autism. It should also be noted that the DSM-IV-TR criteria were written primarily with older children (3 to 5 years old) in mind, because at the time these children made up the majority of research participants and a high proportion of diagnostic referrals (Volkmar et al., 1994). Despite the inappropriateness of these items for very young children, the DSM-IV-TR diagnostic criteria (prior to the advent of DSM-5) remain the standard against which all other diagnostic tools are compared.

DSM-5. Because of the difficulty of distinguishing between AD and PDD-NOS reliably as compared to the differentiation of AD from non-autism spectrum diagnoses, the current diagnostic system, *Diagnostic and Statistical Manual of Mental Disorders: 5th Edition* (*DSM-5;* American Psychiatric Association, 2013) proposed to subsume all the sub-categories of AD, PDD-NOS, AS and childhood disintegrative disorder (CCD) into one category –ASD. In DSM-5, ASD is characterised by impairments in two domains: (1) social communication, and (2) presence of restricted and repetitive behaviours and interests. Although CDD will be integrated into the collective term ASD in the revised DSM, due to the rarity and extreme presentation of this disorder, the present thesis will not involve persons suspected of having this more debilitating form of the disorder.

The rationale for collapsing all the sub-categories of AD, PDD-NOS, AS and childhood disintegrative disorder (CCD) into one category was that clinicians and researchers have reported more difficulties with making the finer distinction between ASD subtypes, especially AD and PDD-NOS (Klin et al., 2000; Mordre et al., 2012; Van Daalen et al., 2009). This finding was also supported by Lord et al. (2012) who found that clinical distinctions among categorical diagnostic subtypes of ASD were not reliable, even when clinicians used standardised diagnostic instruments.

There have been concerns expressed about using the DSM-5 (APA, 2013) criteria for ASD because some individuals who would currently meet criteria under the previous DSM-IV-TR (APA, 2000) would no longer meet criteria under the new DSM-5 due to more rigid scoring criteria where an individual must meet all three of the social communication impairments and at least two of the restricted and repetitive behaviours/interests. In contrast, the lowest threshold for a diagnosis of an ASD under DSM-IV-TR criteria required that a child just needs to demonstrate social impairment (without any specific symptom count) and either communication impairment or RRBs or even sub-threshold presentation across all three areas (Kulage, Smaldone, & Cohn, 2014).

The evidence for the concern about the DSM-5 came from studies done by Gibbs, Aldridge, Chandler, Witzlsperger, and Smith (2012), Matson, Belva, Horovitz, and Bamburg (2012), Matson, Kozlowski, Hattier, Horovitz, and Sipes (2012), McPartland, Reichow, and Volkmar (2012), Worley and Matson (2012), and Young and Rodi (2014) where they compared DSM-IV-TR and DSM-5 criteria for autism/autism spectrum disorders (ASDs). The findings indicated that according to the proposed algorithm, 23–45 % of children,

adolescents and adults classified with ASDs according to DSM-IV-TR criteria, will not meet DSM-5 criteria for ASD. Not all studies concur with these concerns about the DSM-5, however, with Huerta, Bishop, Duncan, Hus and Lord (2012) indicating that the most children (91%) with DSM-IV-TR ASD diagnoses would still be diagnosed as ASD under the proposed DSM-5 criteria. Similarly, Frazier et al. (2012) found similar sensitivity and improved specificity for proposed DSM-5 criteria compared to the DSM-IV-TR.

According to the DSM-5 manual (APA, 2013), children with existing DSM-IV-TR diagnosis of AD, AS or PDD-NOS should be given the diagnosis of ASD. However, this is not assumed in the studies reported in this thesis. Because of the wide variability (23% - 91%) reported in the earlier section in establishing whether children with DSM-IV-TR ASD diagnoses would still be diagnosed as ASD under the proposed DSM-5 criteria, I decided to apply the DSM-5 criteria to the present and archival cases in the sample in my studies to determine whether the participants could still meet DSM-5 criteria.

Screening of ASD

As there is considerable evidence that ASD is a neurodevelopmental disorder with a strong genetic component, there is widespread hope that identifying valid biological markers for ASD will substantially advance research and be readily translated into clinical applications into early identification of ASD (Anderson, 2014). However, the identification of ASD biomarkers has so far proved to be difficult, partly because definitions of the condition itself have changed considerably over time and are still developing and several proposed biomarkers were found not to be universal, and none has indicated the presence of ASD in a majority of cases (poor sensitivity) (Walsh, Elsabbagh, Bolton, &

Singh, 2011). In addition, the measurement of biomarkers can be costly, labour-intensive and needs a high level of technical expertise, thus restricting the possibility of their application in most clinical settings. Similar concerns are also raised regarding the clinical application of reported genetic tests (Kong et al., 2012), neuroimaging tests (Wang, Chen, & Fushing, 2012) and eye-tracking measures (Jones & Klin, 2013). Until these measures can be utilised clinically, paediatricians and clinicians must continue to rely on behavioural, developmental, and historical information to identify ASD.

The American Academy of Paediatrics (AAP, 2006) recommend a system of universal developmental surveillance which is a flexible, longitudinal, continuous, and cumulative process whereby knowledgeable health care professionals identify children who may have developmental problems. In Australia, a large scale developmental surveillance study, the Social Attention and Communication Study (SACS; Barbaro & Dissanayake, 2010; Barbaro, Ridgway, Dissanayake, 2011) had been carried out and aimed to identify key markers of autism in 12- to 24-month-old children. The SACS is a semi-structured play-based assessment that lists a series of social and communicative behaviours together with a list of behaviours of concern related to autism. The clinician scores on the presence of these behaviours and arrives at a 'not at risk' or 'at risk' outcome. In the study, maternal and child health nurses were educated about the signs of ASD in children under 2 years of age. Data obtained from that study indicated that ASD could be identified during children's second year of life through routine developmental surveillance within the Victorian Maternal and Child Health Service. The SACS was designed as an ongoing primary screening tool to be used by community nurses; it has a positive

predictive value of 81% (Barbaro & Dissanayake). It should be noted that the SACS was not designed for children older than 24 months. The authors recommended that infant and toddler monitoring for ASD should become standard practice among all primary health care professionals.

However, there are barriers to developmental surveillance achieving its potential. Surveillance is expensive and requires the training of staff who then engages with the children to conduct the screening. Other barriers include time constraints and difficulties in accessing high quality and affordable primary healthcare for children (Australian Institute of Health and Welfare, 2009).

Surveillance and screening are different in certain aspects. For instance, surveillance is the ongoing and systematic collection of data relevant to the identification of a disorder over time by an integrated health system while screening is the prospective identification of unrecognised disorder by the application of specific tests or examinations (Baird et al., 2001). Yet, at the same time, they are related activities involving the detection of disorders so as to prevent or improve the condition (Baird et al.). Any developmental issues identified through surveillance should be addressed by conducting a structured screening for developmental delays or ASD, or both.

Over the past two decades there has been an increased focus on development of ASD screening instruments. There are a variety of screening instruments available to help clinicians determine the presence of an ASD, ranging from parent checklists to structured interviews to observational tools. A sound ASD screening instrument can provide valuable sources of information about a child who may be at risk of developing ASD and can also help clinicians to make more informed judgements about further referral and diagnostic

services. At this point of time, early identification relies crucially on the availability of standardised and validated screening tools to facilitate clinical judgement, particularly when clinical expertise may be lacking.

I begin my discussion of screening tools with a consideration of the ways in which such tools are typically evaluated. There are two levels of screening for ASD. Level 1 screeners or tools are designed to identify children at risk for ASD in an unselected or low risk population and most likely to be used by primary care practitioners. As such, they should be brief, low-cost and easy to use (Barton, Dumont-Mathieu, & Fein, 2012). Level 2 screeners involve the identification of children at risk for ASD from a population of children referred for developmental concerns, such as general developmental or language delays (Stone, Coonrod, Turner, & Pozdol, 2004). Thus they are designed to be used with a population known to be at risk, and they typically require much more time to administer and interpret (Barton et al.).

Researchers have identified several parameters along which screening tools may be compared (Glascoe, 2005). The first is sensitivity, which is the percentage of true cases correctly identified by a screen; rates of at least 70–80% are the accepted standard. Sensitivity should be high so that children with the disorder will not be missed and the measure does not falsely reassure parents that their children are not at risk (Charman & Baron-Cohen, 2006). A second important criterion is specificity, which is the percentage of non-cases correctly identified; this value should be close to 80% or higher. The sensitivity and specificity of a screening measure are determined by comparing the results of the screener (i.e., risk or no risk) with the diagnostic gold standard for the disorder (Riegelman & Hirsch, 1989). Ideally, a test should have high sensitivity

and high specificity. However, sensitivity and specificity are generally inversely proportional, meaning that as the sensitivity increases, the specificity decreases and vice versa. For example, we can improve the sensitivity of a measure by lowering the cutoff score so that we increase the likelihood of detecting those with the disorder. However, this lower threshold also makes it easier to misidentify those without the disorder as being at risk, which results in a lower specificity (Aylward, 1997). In developing a screening measure, the goal is to identify a cutoff score where we maximise both the sensitivity and specificity (Volkmar, Paul, Rogers, & Pelphrey, 2014).

In most cases, more importance is generally given to sensitivity than specificity in screening instruments (Stone et al., 2004) while specificity is relatively more important than sensitivity in diagnostic instruments (Charman & Gotham, 2013). The main aim of screening is to detect the maximum number of children with the disorder. Therefore, the threshold for identification may be set low, which will lead to the identification of more children with the disorder (high sensitivity), as well as resulting in a significant number of false positives (low specificity). Generally, in the case of screening for ASD, there is greater risk in missing children than in pursuing evaluation of a child who does not have the disorder (Barton et al., 2012). In addition, research suggests that children who falsely screen positive for ASD at 18 months are often at risk of other development disorders (Pandey et al., 2008; Pierce et al., 2011) and, therefore, further assessment may be warranted.

Sensitivity and specificity measure the accuracy (proportion of individuals that the test will correctly classify as either at risk or not at risk) of the screening measure. However, it is also important to consider the proportion

of individuals identified as being at risk (or not at risk) by the screening measure who actually have (or do not have) the disorder (Riegelman & Hirsch, 1989). Hence the third criterion is positive predictive value (PPV), which measures the proportion of children who screen positive (i.e., at risk) who actually have ASD. It may be understood as the inverse of the false positive rate. Finally, some researchers consider negative predictive value (NPV) or the proportion of children who screen negative (i.e., not at risk) who do not have ASD. The predictive power of the test (PPV and NPV) provides information regarding the probability that the individual has or does not have the disorder. Although the PPV is often considered to be the most useful information for the clinician (Camp, 2006), the predictive value of the screening measure will change as a function of the prevalence rate of a disorder in the population under study, given the known sensitivity and specificity of the instruments (Clark & Harrington, 1999). Since that rate will differ across populations (e.g., high risk versus unselected samples,) and because there is considerable variation in the manner in which PPV and NPV are measured and reported for different instruments, it can be difficult to interpret these indices (Barton et al., 2012). These key indices are summarised in Table 1.

As the focus of this thesis is on the ADEC (Young, 2007), which is primarily designated as a Level 2 tool, the following sections will provide a review of other existing Level 2 screening tools available. A broad overview of some of the available Level 1 screening tools can be found in Barton et al. (2012)'s, García-Primo et al. (2014)'s and Johnson et al. (2007)'s papers, and thus will not be covered in this thesis.

Table 1

Summary of Key Indices of Screening Test Accuracy

Sensitivity

The ability of the screening test to correctly identify a high proportion of the

individuals suspected of having the disorder.

Sensitivity = (true positive) / (true positive + false negative)

Specificity

The ability of the screening test to correctly identify those individuals who do

not have the disorder.

Specificity = (true negative) / (true negative + false positive)

Positive predictive value (PPV)

The proportion of individuals identified by the screener who actually have the

disorder.

PPV: = (true positive) / (true positive + false positive)

Negative predictive value (NPV)

The proportion of individuals with negative screening results who do not have

the disorder.

NPV: = (true negative) / (false negative + true negative)

Example 1

Screening tool X, sensitivity = 0.80, specificity = 0.85, prevalence *1000* per

10,000.

For every 10,000 children screened with this screening tool, we would expect

800 with a disorder to be correctly identified as such, 1350 false positives, 200

false negatives and 7650 true negatives. PPV = 0.37, NPV = 0.97

Example 2

Screening tool X, sensitivity = 0.80, specificity = 0.85, prevalence *500* per 10,000.

For every 10,000 children screened with this screening tool, we would expect 400 with a disorder to be correctly identified as such, 1425 false positives, 100 false negatives and 8075 true negatives. PPV = 0.22, NPV = 0.99

Review of Level 2 Screening Tools

While Level 1 tools are useful for general population screening, they are not designed to differentiate ASD from non-ASD conditions such as developmental or speech delay (i.e., Level 2 screeners). Differentiating between children with and without ASD can be challenging. For instance, prevalence studies have reported rates of comorbidity of intellectual disability (or developmental delay) and AD at approximately 50% (Centers for Disease Control and Prevention; CDC, 2002). In addition, individuals with developmental delay often present with autistic behaviours such as poor social connectedness, delayed or absent speech, and stereotyped movements (Matson & Shoemaker, 2009; Wilkins & Matson, 2009). While intellectual disability has never been a component of the diagnostic criteria for AD, the authors of the DSM-IV-TR noted an associated diagnosis of intellectual disability ranging from mild to profound in 70-75% of children (APA, 2000). There is a critical need for further research to develop valid screening tools that can identify the presence of an ASD in young children referred for developmental difficulties.

In this section, I focus on some of the commonly used Level 2 screening tools that are designed specifically for young children below the age of three

years. It should be noted that there are other Level 2 screening tools such as the

Social Communication Questionnaire (SCQ; Rutter, Bailey & Lord, 2003) and

the Social Responsiveness Scale (SRS; Constantino & Gruber, 2005), but they

are designed for older children (four years and above) and thus are not covered

in the following section. A review of these Level 2 tools can be found in Norris

and Lecavalier (2010)'s paper. However, it should be noted that there is now a

preschool version of the Social Responsiveness Scale (Social Responsiveness

Scale–Second Edition, SRS-2; Constantino, & Gruber, 2012), which is valid for

children 30 months and older. The screening tools of interest here include the

Baby and Infant Screen for Children with autism Traits (BISCUIT), specifically

the BISCUIT-Part 1 (Matson, Boisjoli, & Wilkins, 2007; Matson, Wilkins, &

Fodstad, 2011), the Childhood Autism Rating Scale (CARS; Schopler et al.,

1988), the Screening Tool for Autism in Two-Year- Olds (STAT; Stone,

Coonrod, & Ousley, 2000; Stone, Coonrod, Turner, & Pozdol, 2004) and the

SRS-2. The following section provides a review of some of these tools (refer to

Table 2 for an overview of some of the screening instruments and also Johnson

et al., 2007).

Baby and Infant Screen for Children with aUtIsm Traits

(BISCUIT). A relatively new autism-specific parent report screening tool, Baby

and Infant Screen for Children with aUtIsm Traits (BISCUIT), specifically the

BISCUIT-Part 1 (Matson et al., 2007; Matson et al., 2011) was developed to aid

in the assessment of autistic symptomatology and associated features in young

children between the ages of 17–37 months. The BISCUIT-Part 1 contains 62

items designed to aid in the diagnosis of autism and PDD-NOS. Items are rated

through a parent-interview format along a 3-point Likert-type scale with respect

to how the child being assessed compares to a typically developing peer as either: 0 (not different; no impairment), 1 (somewhat different; mild impairment), or 2 (very different; severe impairment). Test administration time of the BISCUIT-Part 1 is approximately 20–30 min, which makes it a relatively time-intensive tool to use. What makes the BISCUIT-Part 1 different from other early autism instruments in that it can differentiate between autism and PDD-NOS in children already identified as being ''at risk'' in the context of a comprehensive evaluation (Matson et al., 2011). Psychometric studies have been published, demonstrating that the BISCUIT-Part 1 has excellent reliability and validity (Matson et al., 2009a; Matson et al., 2009b; Matson et al., 2011).

Childhood Autism Rating Scale (CARS). One widely used rating scale for the detection and diagnosis of autism is the 15-item CARS (Schopler et al., 1988). CARS aids in evaluating the child's body movements, adaptation to change, listening response, verbal communication and relationship to people. It is suitable for use with children over 2 years of age. The examiner observes the child and also obtains relevant information from the parents. The child's behaviour is rated on a 7-point continuum (including midpoints) based on deviation from the typical behaviour of children of the same age, ranging from normal behaviour (1) to severely abnormal behaviour (4). All the item scores are then summed up to give a total score. Total scores of 30 or above are considered to be in the autism range. The psychometric properties of the CARS have been well documented (Nordin, Gillberg, & Nydeˊn, 1998; Perry & Freeman, 1996; Schopler et al., 1988; Tachimori, Osada, & Kurita, 2003). A relatively recent study was even done to look at using the CARS to differentiate children with autism and PDD-NOS (Chlebowski, Green, Barton, & Fein, 2010).

Table 2

Overview of the Level 2 Screening Instruments for Young Children with ASD

Screening Tool	Admin time (min)	Admin age (months)	Admin method	Items	Sen	Spe	Strengths	Limitations
BISCUIT; Matson et al., 2007	20-30	17-37	Parent rated	62	0.84- 0.94	0.67- 0.80	- Can be used for children below 24 months old - Able to differentiate between autism and PDD-NOS	- The BISCUIT takes relatively longer time to administer than other Level 2 screeners - Possible reporting bias in parent rated tools
CARS; Schopler et al., 1998/ CARS2; Schopler et al., 2010	15-20	>24	Clinician rated	15	0.92- 0.98	0.85	- Well documented psychometric properties - Able to differentiate between autism and PDD-NOS - Revised version now to identify those on the "high-functioning" autism spectrum	- Appears to over-diagnose 2-years-old children as having autism - Limited psychometric data on the revised version's usefulness in identifying young children with suspected ASD at this time

STAT;	20	24-36	Clinician	12	0.83	0.86	- Interactive measure which provides a standard set of items or activities that afford direct observation of key behaviours	- Not designed to detect children with PDD-NOS
Stone et al.,			rated					- Requires more training and expertise than parent questionnaires
2000, Stone								
et al., 2008							- Possible usage for children below 24 months	
SRS-2	15-20	30-54	Parent	65	-	-	- Discriminates both within the autism spectrum and between ASD and other disorders	- No predictive data were reported for the Preschool Form
(preschool			rated					- Not meant for children below 30 months old
version);							- A standardisation sample that is more representative of the U.S. population	
Constantino,								- Possible reporting bias in parent rated tools
& Gruber,								
2012								

Note: *BISCUIT* Baby and Infant Screen for Children with aUtIsm Traits, *CARS* Childhood Autism Rating Scales,
STAT Screening Test for Autism in Two-year olds, *SRS-II* Social Responsiveness Scale–Second Edition, *Admin*
Administration, *Sen* Sensitivity, *Spe* Specificity

Although it is highly sensitive, the CARS appears to over-diagnose young children as having autism. Lord (1995) found that the CARS consistently classified children with developmental delay and non-autism as having autism in a sample of 2-year-olds referred for possible autism. Lord (1995) reported that a CARS cutoff score of 30 correctly classified 61.5% of the non-autistic children and 93.7% of the children with autism; however, increasing the CARS autism cutoff to 32 improved classification and accurately classified 84.6% of the non-autistic children, while still correctly classifying 93.7% of the children with autism. Perry, Condillac, Freeman, Dunn-Geier, and Belair (2005) also found that CARS may over-identify older non-verbal children with mental ages below 18 months.

There is now a revised CARS-Second Edition (CARS2; Schopler, Bourgondien, Wellman, & Love, 2010) which helps to identify not only those low-functioning individuals with classic autistic symptoms (Standard Version; CARS2-ST) but also those on the "high-functioning" end of the autism spectrum—that is, those with average or higher IQ scores, normal language abilities, and milder autistic symptoms (High-functioning Version; CARS2-HF). Both the CARS2-ST and CARS2-HF provide cutoff score values intended to inform examiners of further need for evaluation of the presence of ASD. In the development sample for the CARS2-HF, a cutoff score of 28 (with sensitivity of .91 and specificity of .87) was recommended to identify the presence of ASD in the "high-functioning" population (Schopler et al., 2010). Reliability was appropriate for informing diagnosis and research while validity was adequately established and indicated that interpretation of scores from the CARS2 is accurate across settings, informants, and age groups (Vaughan, 2011). However,

there is limited psychometric data bearing on its usefulness in identifying young children with suspected ASD at this time.

Screening Tool for Autism in Two-Year- Olds (STAT). The STAT (Stone et al., 2000; Stone et al., 2004) is unique among the existing screeners in that it is the only Level 2 measure comprised of interactive items. The advantage of an interactive measure is its provision of a standard set of items or activities that afford direct observation of key behaviours. The STAT is intended to be an ASD-specific screener, used in clinics or other specialty centres to identify children at risk for ASD. It is not intended to be a diagnostic measure, and it is designed for use with children 24–36 months of age. Items were selected for inclusion on the STAT based on their effectiveness in differentiating 2-year-old children with autism from developmentally matched children with developmental delay and non-autism. The STAT consists of 12 items organised into four domains: play (two items), requesting (two items), directing attention (four items), and motor imitation (four items). The four domain scores are weighted equally to derive a total score and compared to an empirically derived cutoff score. Initial research with the STAT has revealed strong sensitivity, specificity, and predictive values (Stone et al., 2000). Its interactive format also allows for greater standardisation and assessment of qualitative differences between autism and non-autism clinical groups that may be difficult to ascertain in a questionnaire format and it generates rich observational data that can help guide initial intervention planning. There is also suggestion that the STAT can be used for children under 24 months of age (Stone, McMahon, & Henderson, 2008), though results indicated that the use of a higher cutoff score was required to obtain adequate sensitivity (95%) and specificity (73%) for children 12 to 23

months. The potential limitations of the STAT are that it (a) is not designed to detect children with PDD-NOS and (b) requires more training and expertise than parent questionnaires.

Social Responsiveness Scale–Second Edition (SRS-2). The SRS-2 (Constantino, & Gruber, 2012) is a 65-item, ordinally scaled (1 = "not true" to 4 = "almost always true") quantitative assessment of the severity of autism traits. The scale can be completed by multiple raters who have at least 1 month of experience with the rated individual and it takes approximately 15 to 20 min to complete (Bruni, 2014). Predictive validity data (i.e., sensitivity and specificity) was only available for the School-Age Form (4 to 18 years old), with a sensitivity value of .92, and specificity value of .92. However, no predictive data was reported for the Preschool Form (Bruni) and at this point of time, there is limited psychometric data bearing on its usefulness in identifying young children with suspected ASD.

In summary, identification of ASD made at a young age (below 3 years old) is generally possible, reliable and stable throughout the preschool years. The need for development of screening tools especially for young toddlers arises from the awareness that early identification will lead to early intervention. I have also reviewed some of the common Level 2 screening tools used in identifying young children with suspected ASD and their usefulness and limitations. However, there are limited Level 2 screening tools that are designed and validated for young children and also have an interactive component in the administration. The next section will describe the Autism Detection in Early Childhood (ADEC, Young, 1997), which is the main focus of this thesis, in greater detail.

The Need for Autism Detection in Early Childhood (ADEC)

While some of the Level 2 screening tools reviewed in the earlier section have excellent test-retest reliability, internal consistency and concurrent validity, they have a number of limitations. For example, only the BISCUIT can be used with young children from the age of 17 months of age. Other commonly used tests such as the STAT and CARS are only suitable for children beyond 24 months. Although some evidence points to the clinical utility of the STAT for children under 24 months of age (Stone et al., 2008), the use of a higher cutoff score was required to obtain adequate sensitivity and specificity for children 12 to 23 months. Furthermore, the CARS over-diagnosed young children in a sample of 2-year-olds referred for possible AD (Lord, 1995) and hence may not be a suitable screening tool for use with young children.

Both the BISCUIT and SRS-2 are parent report measures and are possibly limited because parents may be unfamiliar with both the characteristics of ASD and what might be considered atypical behaviour. Parental report measures are often less reliable than direct observation because parents may use compensatory strategies in an effort to engage their children more successfully in social interactions and play (Baranek, 1999). Hence, follow-up interviews are often required to improve the accuracy of the screening measure (Eaves, Wingert, Ho, & Mickelson, 2006). Young et al. (2003) found that almost three quarters of children who screened positive for AD using an AD-specific screening instrument did not elicit corresponding developmental concerns from their parents as measured by a standardised general developmental questionnaire. This indicates that parent report measures may not be reliable when used with very young children with ASD (Barton et al., 2012; Wiggins,

Bakeman, Adamson, & Robins, 2007). In addition, it has been suggested that
interactive screening tools may allow the clinician to experience the child's
social and communicative behaviours firsthand, which in turn can help to
inform clinical judgment, and also can help to identify the child's strengths and
weaknesses through the activities (Volkmar et al., 2014). Therefore, what is
needed is a brief, valid and reliable observation/interactive measure.

The Autism Detection in Early Childhood (ADEC; Young, 2007) has the
potential to address some of these shortcomings of the Level 2 screening tools
noted above. The ADEC is a 16-item observation checklist developed to identify
AD in young children between the ages of 12 to 36 months and focuses on
preverbal behaviours which are not dependent on receptive language abilities.
The assessor interacts with the child with the aim of eliciting 16
developmentally appropriate behaviours. The specific behaviours that are
observed during the administration of the ADEC are: (1) response to name, (2)
imitation, (3) ritualistic play, (4) joint attention and social referencing, (5) eye
contact, (6) functional play, (7) pretend play, (8) reciprocity of smile, (9)
reaction to common sounds, (10) gaze monitoring, (11) following verbal
commands, (12) delayed language, (13) anticipation of social advances, (14)
nestling, (15) use of gestures, and (16) task switching. Response scores for each
item range from 0 (appropriate) to 2 (inappropriate), with a possible maximum
score of 32 (refer to Appendix A for details of all ADEC items and scoring
protocols and Appendix B for the score sheet). Based on sensitivity and
specificity data provided in the manual, a score of 0-10 indicates a low risk for
AD, 11-13 a moderate risk, 14-19 a high risk, and >19 a very high risk.

The ADEC is appealing for several reasons. First, the ADEC comprised of 16 discrete behaviours thought to reflect the core deficits of ASD which can be identified in very young children. These behaviours were identified from retrospective parental reports (Young et al., 2003) and video analysis (Clifford, Young, & Williamson, 2007); they have been clearly operationalised and thus can be measured reliably. By clearly operationalising each item of behaviour, and providing examples and non-examples of appropriate responses (in the manual), subjective interpretation on the part of the assessor is minimised and more precise (and reliable) scoring can be obtained. For example, *social response* has been operationalised as whether a child responds to his/her name when called by the examiner over 5 trials. If the child responds in the first or second attempt it is scored a 0, in 3-5 attempts it is scored a one, and a two is given if the child does not respond to his/her name in any of the five attempts (for more examples of scoring, refer to Appendix A).

The ADEC's scoring approach also helps to capture the finer nuances of behaviour of children with ASD, allowing for the possibility that children at risk of ASD may present with a reduced rate of key behaviours as well as a wide spectrum of autistic behaviours (e.g., Allison et al., 2008; Constantino et al., 2006). For example, a score of 1 can be given to children who did not perform the action (e.g., respond to his/her name when called) when required by the examiner but displayed the behaviour spontaneously at other times during the testing session. In other words, the child still receives partial credit rather than 'failing' the item. Items in other screening tools (e.g., the STAT) are scored merely as pass or fail. This indicates only that the child did (credit given) or did not perform (credit not given) the required action at the particular test prompt

given, thereby potentially limiting those tools' usefulness in capturing the

heterogeneity and spontaneity of behaviours typically observed in children with

ASD. Unlike the STAT, CARS and the BISCUIT, the ADEC is specifically

designed to detect AD in children as young as from 12 months old. Finally, the

ADEC can be administered in a relatively short amount of time (10-15 minutes)

by an administrator with limited clinical training. In contrast, the STAT requires

about 20 minutes for administration while the CARS and the BISCUIT take

about 20 to 30 minutes.

The ADEC manual (Young, 2007) provides promising data on internal

consistency (Cronbach's α ranging from .85 to .93), test-retest reliability (r =

.83) and inter-rater reliability (intra-class correlation coefficient, ICC = .83). A

Spanish version of the ADEC (ADEC-SP) has also been validated as a

promising screening tool for use in a Hispanic sample of children aged 15-73

months (Hedley, Young, Juarez-Gallegos, & Marcin-Salazar, 2010).

Study Objectives

The purpose of this thesis is to assess the psychometric properties of a

relatively new observation measure, the ADEC. Given the potential of the

ADEC to become a useful component of a comprehensive assessment for ASD,

the current thesis is a relevant first step in establishing the strength of this

measure through the three studies conducted.

Study 1. Study 1 provides a comprehensive psychometric validation of

the ADEC as a screening tool for ASD. Because the ADEC was originally

designed to be used for young children aged between 12 to 36 months old, this

study examined how well the ADEC classifies young children with a DSM-IV-

TR diagnosis of AD within this specified age group when compared to "gold

standard" ASD diagnostic measures such as the ADOS (Lord et al., 2000), the ADI-R (Le Couteur et al., 2003) and clinical judgment based on DSM-IV-TR (APA, 2000).

The screening properties of the ADEC were also be examined using receiver operating characteristic (ROC) curve analyses to provide data on both sensitivity and specificity associated with different cutoff scores, and to determine the optimal cutoff score to maximise both. As well as establishing whether the ADEC permits reliable and valid test scoring, key objectives were to determine whether it could reliably identify AD among children in the general population and children without ASD. Besides examining diagnostic validity, other aspects of validity such as construct validity (using exploratory factor analysis) and concurrent validity (in relation to ADOS, ADI-R and DSM-IV-TR criteria) were examined in this study.

Study 2. Currently, there is little information available about the validity of Level 2 screening tools in predicting long term outcomes such as diagnostic classification. Establishing the predictive validity of ASD screening tools is important, as it allows clinicians to use information obtained in the administration of the screening test to determine (a) how likely it is that these children will continue to meet diagnostic criteria for ASD with age and, more importantly, (b) the likely severity of this disorder.

In Study 2, I compared the predictive validity data of the ADEC against a well-established screening tool, the CARS (Schopler et al., 1998), in relation to diagnostic classifications, symptom severity and functioning level at 2 and 6 years following initial assessment. To date, no studies have provided this comprehensive evaluation (and comparison) of the predictive validity of Level 2

ASD screening instruments on long term outcome measures in children with ASD.

I am aware of only one other study that has examined the predictive validity of a screening instrument, the Parent Observation of Early Markers Scale (POEMS; Feldman et al., 2012). Preliminary results were promising with respect to diagnostic classifications obtained one to two years after initial assessment. The POEMS was, however, designed as a parent report measure to monitor the behavioural development of infants at risk for ASD and there appears to be no clinician-administered screening tool with predictive validity established for long term outcomes reported.

Study 3. It has been established that screening for ASD is a crucial first step for improving early identification of children at risk of the disorder. Though ASD is considered to be a heterogeneous group of neurodevelopmental disorders, we need to have a better understanding of the variability of symptoms detected among individuals with ASD at different ages, as well as the age at which different behavioural features emerge. Given that Studies 1 and 2 showed the ADEC to be a suitable screening tool for young children, I explored and extended the utility of the ADEC for older children (i.e., older than 36 months) with possible ASD in Study 3.

The reality is that there are some children whose diagnoses do not occur as early as one might hope. According to the CDC, only about 18% of children are diagnosed with an ASD by age 3 years (CDC, 2012), and more than half of children with developmental disabilities (including ASD) are not identified until they enter school, around age 4 years (Sices et al., 2003). In another study by Shattuck et al. (2009), the authors found that children could be identified with

an ASD as late as 6 years old (median 5.7 years). Therefore, having available valid screening tools across a broad range of ages from the very young and school age is required.

It is likely that given the change in presentation of this disorder across time, these screeners may look quite different. There may be certain behavioural markers of ASD that may be present in some younger children which we may not see in older children with ASD, and vice versa. Precise information on the age at which different behavioural features are evident in children evaluated for ASD should highlight opportunities to improve the early detection of ASD.

In the first part of this study, I examined how the frequency and pattern of documented diagnostic features based on the ADEC varied with the age of the children with and without ASD. Specifically, I examined the age groups 12-23 months, 24-35 months, 36-47 months, 48-59 months and 60-71 months and visually inspected whether there were any items for which there was a consistently low frequency of typical behaviours (such as responding to their name when called) displayed by children with ASD across the age groups. Failure to exhibit these typical behaviours may indicate pervasive autistic difficulties in children with ASD regardless of their age. I also examined whether there were any items of typical behaviour that increase in reported frequency in children with ASD with age. Such behaviour patterns may indicate that some older children with ASD have learnt strategies to demonstrate these typical behaviours; alternatively, they may simply reflect maturation.

Improving autism screening tools is one goal, but there is also a need to focus at the same time on adoption of these tools. Practitioners have reported a variety of concerns such as the time and training required regarding using

current ASD screening tools in their practice (Barton et al., 2012). Therefore, it is important to develop *brief* screening tools that are practical for use in a busy practice and can be easily integrated into existing procedures, in addition to having strong psychometric features.

Although the ADEC consists of only 16 items and is reported to require only 10-15 minutes to administer, paediatricians argue this may remain prohibitive given their time constraints. In the second part of Study 3, given that I identified behaviours with better diagnostic salience at different ages, I examined whether I could develop a brief age-specific ADEC (BADEC) version for identifying children with possible ASD. I proposed that the brief version would consist of 3-5 items which should take less than 5 minutes to administer and to score. I examined the sensitivity, specificity, positive and negative predictive values associated with the different BADEC cutoff score. I also calculated the internal consistency, diagnostic validity and concurrent validity for the brief version for each age group. The different BADEC versions will provide clinicians with a tool to enable them to make quick decisions about whether to refer patients to specialist diagnostic services for ASD.

Summary

Over the past two decades there has been an increased focus on development of ASD screening instruments. A sound ASD screening instrument can provide valuable sources of information about a child who may be at risk of developing ASD and can also help clinicians to make more informed judgements about further referral and diagnostic services. Consistent with the views of others (e.g., Charman & Gotham, 2013), we need to know more about

a few select tools rather than a little about a lot of tools as is currently the case. This is the focus of the present thesis.

This thesis aimed to determine whether the ADEC could be a quick and suitable screening tool to help clinicians and paediatricians to identify young children presenting with possible ASD in their practice settings. To summarise, Study 1 provided reliability and validity data for a Level 2 ASD screening tool - the ADEC. Study 2 provided data on the efficacy of the ADEC and the CARS in predicting long term outcomes in children with ASD. Study 3 examined how the behaviours (in terms of ADEC items) changed across the age range for children with and without ASD and then derived a brief version that is suitable for identifying children with possible ASD, across the age range from 12 to 71 months.

CHAPTER 2

Study 1[1]: Autism Detection in Early Childhood (ADEC): Reliability and Validity Data for a Level 2 Screening Tool for Autistic Disorder

It has been noted that in the previous chapter, clinicians and paediatricians have to rely on parental report and behavioural observations to help them to identify children with possible ASD. Existing ASD screening tools for young children have focused primarily on screening at the population level (i.e., Level 1). Some of the more popular Level 1 screening tools include the Checklist of Autism in Toddlers (CHAT; Baird et al., 2000), the modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, Barton, & Green, 2001) and the Pervasive Developmental Disorders Screening Test-II (PDDST-II; Siegel, 2004). While these tools are useful for general population screening, they are not designed to differentiate ASD from non-ASD conditions such as developmental or speech delay (i.e., Level 2 tools).

There is a critical need for further research to develop valid screening tools that can identify the presence of an ASD in young children referred for developmental difficulties while not over including children with these other disabilities. The ADEC was developed to address this need.

Study Aims

The aim of the present study was to provide a comprehensive psychometric examination of the ADEC as a screening tool for Autistic Disorder

---

[1] Material presented as part of this study appeared in the *Psychological Assessment* journal as 'Autism Detection in Early Childhood (ADEC): Reliability and Validity Data for a Level 2 Screening Tool for Autistic Disorder' (Nah et al., 2014).

(AD). In the present study, because the ADEC was originally designed to be used for young children aged between 12 to 36 months old, the focus was on examining how well the ADEC classifies young children with AD within this specified age group when compared to "gold standard" ASD diagnostic measures such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000), the Autism Diagnostic Interview-Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003) and clinical judgment based on DSM-IV-TR (APA, 2000). Because the ADEC is designed to be a screening tool, not a diagnostic tool, sensitivity is a priority as it may be more beneficial to over-identify children at risk (since this group of children may also have other developmental disorders) than fail to identify children at risk for AD or other non-ASD developmental disorders who may benefit from additional services. This in no way diminishes the importance of specificity which is also valued as I want to ensure that those cases without the disorder screen negative (i.e., without the disorder, thereby avoiding causing false alarm to parents and costly referral for in-depth assessment). Nevertheless, the screening properties of the ADEC are also examined using receiver operating characteristic (ROC) curve analyses to provide data on both sensitivity and specificity associated with different cutoff scores, and to determine the optimal cutoff score to maximise both.

The ADEC was administered to a sample of children aged between 12 and 36 months referred for developmental concerns – namely, ASD and other developmental disorders – and typically developing children. Diagnostic evaluation and measurement of cognitive-developmental and adaptive functioning level using the Mullen Scales of Early Learning (MSEL; Mullen, 1995) and the Vineland Adaptive Behavior Scales-Second Edition (Vineland-II;

Sparrow, Cicchetti, & Balla, 2005) were also obtained to ascertain that ADEC scores (a) are not simply dependent on developmental level and (b) can identify children with AD rather than developmental delay per se. The developmental assessments were also used to supplement the decision-making of the best estimate clinical (BEC) DSM-IV-TR diagnoses and to differentiate children with other developmental delays from the AD group. ADEC performance was compared across the three groups. As well as establishing whether the ADEC permits reliable and valid test scoring, key objectives were to determine whether it could reliably identify AD among children in the general population and children with non-ASD. Besides examining diagnostic validity, other aspects of validity such as construct validity (using exploratory factor analysis) and concurrent validity (in relation to ADOS, ADI-R and DSM-IV criteria) were examined.

Method

Participants

The participants were recruited from several sources: (a) general advertising in mass media including newspaper, university newsletter and radio and in child-care centres, government and private developmental clinics where children aged between 12 and 36 months were suspected of having communication and/or developmental delay and invited to participate in an autism screening university research study ($N = 127$), and (b) participation in a university-based autism research center in South Australia ($N = 74$). Only 23 participants' data (31.1%) about SES and parental education level were available for those participants involved in the university-based autism research center. Of these 23 families, the primary care-giver was the mother and they

were generally well educated, with 95.6 % having at least some diploma education. The average total family income fell in the $AUD 60,000–$100,000 range. Demographic information was not collected for the batch of participants involved in the autism screening university research study and some participants involved in the university-based autism research center chose not to provide the details. The ethnic background of the sample was predominantly Caucasian ($N =$ 194, 90.2 %). A small subset of participants ($N = 14$) was also recruited from early intervention centres in Singapore. Like Australia, Singapore has an advanced economy, hosts a multi-cultural society and English is the first language. Data collection was spread over a 10-year period from 2003 to 2013 gathering as large sample as possible with the maximum age of 36 months as a cutoff criterion, and the author was involved in the data collection[2] for this study from April 2011 to December 2012. All parents provided informed consent and appropriate ethics approvals were obtained prior to conducting this study.

A best estimate clinical (BEC) DSM-IV-TR diagnosis was made by the author (who was blind to the ADEC scores) using all available information and assessment results, excluding ADEC data, to generate diagnoses independent of the ADEC. BEC diagnoses using DSM-IV-TR or International Classification of Diseases, Tenth Revision (ICD-10; World Health Organization, 1993) criteria were used as they are commonly used by autism researchers (e.g., Chawarska et al., 2007; Lord et al., 2006) to categorise participants by diagnosis. They have been shown to be reliable (Klin et al., 2000) and generally stable over time, even for children under 3 years of age (Moore & Goodson 2003; Stone et al., 1999).

---

[2] My role includes designing the study, recruiting participants, administering various tests such as the ADOS, ADI-R, Vineland and Mullen Scales to the participants, data analysis and write up.

The resulting sample consisted of 70 children (57 male, 13 female) with a BEC DSM-IV-TR diagnosis of AD, 24 children (16 male, 8 female) with PDD-NOS, 57 children (37 male, 20 female) with non-ASD (such as language and developmental delay, Cerebral Palsy, Down Syndrome, multiple disabilities and hearing loss) and 64 typically developing (TD) children (31 male, 33 female), with all participants aged between 12 and 36 months.

Of the participants with ASD and non-ASD, 77.5% ($N = 117$) had an independent confirmatory diagnosis (i.e., a separate diagnosis apart from the author's BEC diagnosis) from either two other independent professionals who were recognized by the state's autism association or other medical professionals such as paediatricians and psychologists. The author was also blind to the results of this confirmatory diagnosis. The inter-rater reliability for DSM-IV-TR diagnosis of AD between the researcher's BEC diagnosis and the independent diagnosis was substantial ($k = .93$, $p < .001$) though not perfect. Analyses of these data show that the errors made were due to the disagreement between AD and PDD-NOS diagnoses. When I combined both the AD and PDD-NOS into an ASD category, inter-rater reliability increased (k = 1.00, $p < .001$).

Materials

ADEC. Details about the ADEC have been described in the Introduction chapter and would not be repeated here. Briefly, the ADEC is a 16-item observation checklist used to identify AD in young children between the ages of 12 to 36 months. Based on sensitivity and specificity data provided in the manual, a total score of 0-10 indicates a low risk for AD, 11-13 a moderate risk, 14-19 a high risk, and >19 a very high risk.

Autism Diagnostic Observation Schedule (ADOS). The ADOS (Lord et al., 2000) is a semi-structured assessment of communication, social interaction and play or imaginative use of materials for individuals suspected of having ASD. The ADOS consists of four modules, each of which is appropriate for children and adults of differing developmental and language levels, ranging from no expressive or receptive language to verbally fluent adults. The ADOS has demonstrated strong psychometric properties, with inter-rater reliability measured by mean exact agreement shown to be over 88% for ADOS Modules 1–4. One difficulty with the ADOS is that, because it attempts to differentiate between broadly defined ASD and a more narrowly defined autism, the diagnostic categories can shift from autism to ASD and vice versa quite frequently when scores are compared across time or raters. Moreover, because the ADOS uses four modules with different tasks and language demands to decrease the effect of language level on diagnosis, a child who experiences gains in language will likely be administered a higher module that has more difficult social and communicative demands. If the child's gains in language are not accompanied by similarly large gains in social and communicative skills, the child could appear more severely autistic when assessed with the more advanced module. On the other hand, a child who shows slow gains in language skill (or no gains at all) will likely not change modules and, as a result, may show a change in diagnostic category, particularly from autism to PDD-NOS, because of the eventual acquisition of very basic skills.

The ADOS's original algorithm was revised in 2007 (Gotham et al., 2007) which helped to increase specificity in classifying non-autism ASD in lower functioning children without any words and the modest gain in specificity

for older children who have not progressed beyond phrase speech. The ADOS

calibrated severity score (Gotham, Pickles, & Lord, 2009) is then derived from

the ADOS's revised algorithm raw total score. This severity metric offers a

method of quantifying ASD severity after controlling for individual

characteristics such as age and verbal IQ.

In the revised algorithms, the original ADOS domains and cutoff values

for Social and Communication items have been collapsed into a single factor

consisting of 10 items that describes social and communication domain items:

the Social Affect factor (SA). In addition, 4 items from a second factor,

restricted, totals repetitive behaviour (RRB), have been included because RRB

domain items may contribute to the diagnosis of autism or ASD, even in the

limited context of the ADOS (Lord et al. 2006). There are two diagnostic cutoff

scores for the combined SA and RRB domain total, one for autism and one for

ASD. In order to reduce ceiling effects in communication items, the revised

algorithms distinguish between ''Some words'' and ''No words'' in Module 1,

on the basis of the item A1 score (overall level of non-echoed language). To

reduce the difference between younger more rapidly developing children and

older children, the revised algorithms distinguish between age younger and older

than 5 years in Module 2 (Gotham et al., 2007).

At the time of data collection, the ADOS-2 (Lord et al., 2012) had not

been released yet and so, I used the revised ADOS algorithms (Gotham et al.,

2007; Oosterling et al., 2010) in this study. It should be noted that the revised

ADOS algorithms were the same as the ADOS-2 algorithms. Items are scored

on a 4-point scale, with 0 indicating "no abnormality of type specified" and 3

indicating "moderate to severe abnormality." Item scores of 2 and 3 are

collapsed in the algorithms to reduce the impact of individual items. The total revised ADOS algorithm score ranges from 0 to 28. Therefore, Study 1 used the revised algorithm over the original algorithm. In addition, the revised ADOS algorithm could be used to derive an autism severity score (Gotham et al., 2009). Study 1 assessed the relationship between the ADEC score and this autism severity score, besides examining the ADOS revised algorithm score.

Autism Diagnostic Interview-Revised (ADI-R). The ADI-R (Le Couteur et al., 2003) is a standardised, semi-structured clinical interview for caregivers of children and adults. The ADI-R is appropriate for children and adults with mental ages from about 24 months and above. One major criticism of using the ADI-R to diagnose toddlers is that it only diagnoses autism, and not PDD-NOS. Since, to diagnose autism (based on the DSM-IV-TR), a child must show deficits in all three areas (socialization, communication, and stereotyped/ repetitive behaviours), and many young children only show problems in the first two areas, using the ADI-R may fail to diagnose such children with an ASD (Charman & Baird, 2002; Cox et al., 1999; Lord, 1995; Stone et al., 1999; Ventola et al., 2006). In addition, many children with severe global developmental delay may meet diagnostic criteria for autism on this measure because they are engaging in a number of repetitive mannerisms, even at a very young age (Lord, 1995). Probably for all of these reasons, a study considering diagnoses of 2-year olds (Ventola et al.) found that clinical judgment, ADOS, and CARS agreed with each other in determining whether a child has autism or PDD-NOS, but this was not the case for the ADI-R. The children classified with PDD-NOS by the other measures were not classified with autism using the ADI-R because they did not display enough repetitive and restricted behaviours.

Cox et al. (1999) found that early diagnosis based on the ADI-R showed good specificity, in that all children diagnosed with autism at 20 months met diagnostic criteria at 42 months. However, it showed poor sensitivity in detecting autism at 20 months, in that a high proportion of children later found to have ASD were missed at 20 months. This issue is less of a problem with other screening/diagnostic measures, such as the CARS and the ADOS, because these measures allow for a diagnosis of PDD-NOS that does not require repetitive and stereotyped behaviours. This may help increase the negative predictive power of these measures for children in this age range as the presence of repetitive behaviours does not differentiate children with ASD and children with other developmental delays (Baranek, 1999; Lord et al., 1993; Stone & Hogan, 1993).

Most items are scored from zero (no impairment with respect to the behavioural definition for each item) to three (severe impairment for the individual and their family), relying on the interviewer to make judgement on the child's behaviour based on the recall of information from parents/carers. Scores are transformed following the protocol in the manual (e.g. 3's become 2's). The total ADI-R score (i.e., the sum of communication, social interaction, and patterns of behaviour scores) ranges from 0 to 68. A classification of AD is given when scores in the three content areas of communication, social interaction, and patterns of behaviour on the ADI-R meet or exceed the specified cutoffs, and onset of the disorder is evident by 36 months of age. The algorithm specifies a minimum score in each area to yield a diagnosis of AD as described in the DSM-IV-TR. Study 1 also assessed the relationship between the ADEC

score and the total ADI-R score (i.e., the sum of communication, social interaction, and patterns of behaviour scores).

MSEL. The Mullen Scale of Early Learning (MSEL; Mullen, 1995) is a developmental test intended for use in children aged 0 to 68 months and yields an Early Learning Composite score ($M = 100$, $SD = 15$). It is a standardised test that measures ability in four domains: fine motor, receptive language, expressive language, and visual reception. The MSEL demonstrate satisfactory internal consistency ($\alpha = 0.75$–0.83), test–retest reliability (0.71–0.96), and inter-rater reliability (0.91–0.99) (Mullen, 1995). Concurrent validity is indicated by high correlations between the MSEL Early Learning Composite and Bayley Mental Development Index ($r = .70$), the MSEL fine motor scale and the Peabody fine motor scale ($r = .65$-$.82$), the MSEL receptive language scale and Preschool Language Assessment (PLA) auditory comprehension scale ($r = .85$), and the MSEL expressive language scale and the PLA verbal ability scale ($r = .80$) (Mullen). However, comparison of performance between the AD and non-ASD groups in this study is better illustrated using age-equivalent scores, because many T scores across the assessments (in mainly the AD cases) were three or more standard deviations below the mean (i.e., T = minimum score of 20). Non-verbal IQ (NVIQ) was calculated by using the age equivalent scores from the visual reception and fine motor scales and divided by the child's chronological age. There were two participants that were administered the Bayley Scales of Infant and Toddler Development, Third Edition Screening Test (Bayley-III Screener; Bayley, 2005) instead of the MSEL. As the Bayley-III Screener does not provide any age equivalent scores, the scores were then pro-rated to give a full Bayley-III score (e.g., the expressive communication score of the Bayley-III

Screener was multiplied by 2 to get the equivalent score in the full Bayley-III).

The VIQ (using age equivalent scores from the receptive and expressive

communication scales) and the NVIQ (using age equivalent scores from the

cognitive, gross and fine motor scales) were also calculated using the same

procedure as for MSEL. While I attempted to get a measure of verbal IQ, testing

was considered invalid and the data were not used. This is consistent with other

attempts to assess verbal IQ in young children with ASD where the children

typically present with relatively spared non-verbal skills (e.g., visual

discrimination and memory, visual-motor coordination) and more significantly

impaired verbal abilities (Chawarska, Klin, Paul, Macari, & Volkmar, 2009;

Steiner, Goldsmith, Snow, & Chawarska, 2012), with the latter thought to

reflect autism severity rather than IQ per se (Munson et al., 2008). In addition,

children with ASD often present significant discrepancies between their verbal

and non-verbal IQ (Joseph, Tager-Flusberg, & Lord, 2002) and, thus, an overall

IQ score may also not be meaningful.

Vineland. The Vineland Adaptive Behavior Scales (Vineland; Sparrow,

Balla, & Cicchetti, 1984 and Vineland-II; Sparrow et al., 2005) provide a

measure of adaptive behaviour from birth to adulthood and are designed to

obtain information about skills a person consistently demonstrates to adapt and

function in everyday routines within their environment. This measure yields a

standard score in four domains—Communication, Daily Living, Social, and

Motor; an Adaptive Behavior Composite (ABC) provides an index of adaptive

behaviour/developmental maturity. A standard score of 100 (SD = 15) reflects

mean performance at any given age. Reliability data include internal consistency

(0.70-0.97), test-retest (0.70-0.90), and inter-rater (0.70-0.80) for the

parent/caregiver interview forms. The manual also reported correlations between the Vineland-II and Vineland Adaptive Behavior Scales (1984)s' adaptive behaviour domains (0.65-0.94); the ABC (0.82-0.91) and the Vineland-II and Adaptive Behavior Assessment System, Second Edition (ABAS-II; Harrison & Oakland, 2003)'s composite scores for children birth through 5 years (0.63).

Procedure

Parents and health care professionals who were concerned that their child or client presented with possible risk of developing an ASD participated in this screening study and the children were assessed with a battery of tests (refer to Table 3) either at the university autism research center or at the participant's home. The ADEC was administered to the participants by (four different) examiners who had university degrees in psychology or psychology-related courses and had received prior training (which consisted of at least reading the manual and viewing the training DVD) on administration and scoring, though none were professionals suitably qualified to diagnose ASD. While the ADEC was being administered by one examiner, a second independently observed and scored the administration to ascertain inter-rater agreement. The ADEC administration was independent of the diagnostic assessment and the administrators were blind to the results of the diagnostic evaluation. Likewise, the diagnostic assessor who administered the ADOS and ADI-R was blind to the ADEC assessment result. Because of the long administration time of the ADI-R, only participants ($N= 21$) who obtained either an ASD or AD classification on the ADOS and/or presented with some autistic traits were then followed up with the ADI-R to confirm the diagnosis of ASD. Diagnoses were based on clinical

judgements using the DSM-IV-TR criteria (APA, 2000) together with the

ADOS and ADI-R information. The diagnostic assessor was fully trained to the

standards of research reliability for ADOS and ADI-R set by the test developer.

Children who were deemed by the clinician, parents and child-care

workers to be typically developing (i.e., no concerns were expressed) were only

administered the ADEC. M-CHAT data were available for 44 of the 64 TD

participants and were examined to confirm the absence of a possible ASD

diagnosis. Only one participant, aged 12 months, failed the M-CHAT, though

clinical judgement deemed that he did not meet diagnostic criteria for ASD or

any developmental difficulties.

Participants returned to the research unit for a second ADEC

administration an average of 54.5 days following their first (range = 12-138

days) to gauge test-retest reliability. Different examiners administered the

ADEC at each point in time to reduce any potential familiarity bias.

Results

There were 12 participants (5%) who had missing data on some of the

ADEC items. These participants' data were retained and pairwise deletion

analysis was used. Sample characteristics (based on BEC DSM-IV-TR

diagnoses) are presented in Table 3. Given the group differences on the non-

verbal IQ variable approached statistical significance, $F(2, 65) = 2.69$, $p = .08$, I

created a subset of participants (AD and non-ASD) matched on age, non-verbal

IQ and VABC for comparison purposes (see Table 3) to ensure any differences

found were not dependent upon IQ.

Reliability

The ADEC's internal consistency was excellent ($\alpha = .91$) and did not improve meaningfully with the removal of any specific item. Inter-rater reliability of the ADEC total score was available for 19.5% of the participants and was high, $r(42) = .98$, $p < .001$. The test-retest reliability of the ADEC total score was also high, $r(62) = .72$, $p < .001$.

Validity

The validity analyses described below demonstrated: (1) construct validity of the ADEC as a one-factor solution (Social Communication), (2) high concurrent validity of the ADEC with the ADOS, ADI-R and DSM-IV-TR classification of AD, and (3) diagnostic validity of the ADEC across the different groups (AD, PDD-NOS and non-ASD).

Construct validity. Construct validity was assessed by conducting a preliminary investigation of the factor structure of the ADEC's 16 items with Principal Axis factoring (PAF) using all 215 participants. I found that, for my sample, skew was -.21 to 1.88 (*SE*= .17) and kurtosis was -1.82 to 2.56 (*SE*= .33), both of which were within the recommended ranges for item skewness (±3) and kurtosis (±7) (West, Flinch, & Curran, 1995). In this study, it was hypothesised that, consistent with the current DSM-IV-TR factor structure, constructs reflecting the core autistic symptoms (i.e., difficulties in social interaction and communication, unusual play, and sensory behaviours) in young children with ASD would emerge.

Table 3

Means and Standard Deviations for Sample Characteristics

	BEC DSM-IV Diagnosis (unmatched sample)				BEC DSM-IV Diagnosis (matched sample)	
	AD	PDD-NOS	Non-ASD	TD	AD	Non-ASD
Measures	($n = 70$)	($n = 24$)	($n = 57$)	($n = 64$)	($n = 39$)	($n = 39$)
Chronological Age (months)	29.4 (5.1)	28.2 (7.0)	24.1 (7.1)	23.5 (6.8)	27.7 (4.9)	25.4 (6.2)
ADEC	19.0 (5.4)	9.9 (6.3)	8.5 (6.1)	2.7 (3.0)	19.2 (5.6)	9.2 (7.0)
ADOS Rev	16.5 (5.3)	12.0 (4.0)	5.9 (3.9)	-	16.4 (5.6)	7.1 (4.8)
ADOS Severity	6.1 (2.3)	4.9 (1.7)	2.4 (1.5)	-	6.3 (2.4)	2.8 (1.5)
ADI-R Total	35.2 (7.6)	26.6 (5.1)	-	-	36.2 (7.7)	-
Non-Verbal IQ	48.6 (10.2)	70.6 (13.0)	63.2 (27.5)	-	49.7 (9.5)	53.1 (30.2)
VABC[a]	62.1 (7.9)	72.7 (7.3)	77.1 (8.1)	-	67.1 (7.0)	71.1 (7.8)

Note: ADOS Rev Autism Diagnostic Observation Schedule – Revised total algorithm score, *ADI-R Total* Autism Diagnostic Interview – Revised total algorithm score
[a] $n = 57$ completed the Vineland Adaptive Behavior Scales; $n = 52$ completed the Vineland Adaptive Behavior Scales 2nd Edition

The Kaiser-Meyer-Olkin measure (KMO = .92) exceeded the recommended value of .6 (Tabachnick & Fidell, 2001), verifying the sampling adequacy for the proposed analysis. Bartlett's Test of Sphericity, χ^2 (120) = 1548.3, $p < .001$, indicated that correlations between items were sufficiently large for PAF. An oblique rotation, direct oblimin (delta = 0), was used due to the likelihood of psychological constructs being highly correlated (Fabrigar,

Wegener, MacCallum, & Strahan, 1999). A combination of methods was used for retention of factors, including the Kaiser's criterion (retention of factors with eigenvalues >1.0), the Cattell's scree test (examination of a plot of the eigenvalues for breaks or discontinuities), parallel analysis, simple structure and interpretability. To determine inclusion in a factor, a score above .40 on a primary loading of items after rotation was used as the cutoff. Correlations between the ADEC items are presented in Table 4, along with ADEC item means and standard deviations.

The PAF revealed the presence of four factors with eigenvalues exceeding 1, explaining 43.6%, 8.3%, 7.9% and 6.3% of the variance, respectively. While the first factor was dominant, the scree plot was slightly ambiguous and showed inflexions that would justify retaining both one and four factors. Next, a parallel analysis was conducted using the software developed by Watkins (2000) to estimate and confirm the number of factors (Zwick & Velicer, 1986). The size of eigenvalues obtained from PAF was compared with those obtained from a randomly generated data set of the same size. Only factors with eigenvalues exceeding the values obtained from the corresponding random data set were retained for further investigation. The results of the parallel analysis showed only one such factor. Consequently, the PAF was re-run to extract a one factor solution: Factor 1 – Social Communication (items 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 16). The internal consistency of Factor 1 was indicated by $\alpha = .92$. The exclusion of any of the above items did not result in a substantial increase in α for this factor (see Table 5).

Table 4

Univariate Summary Statistics and Inter-item Correlations (Pearson in Upper and Polychoric in Lower Part of Matrix) of the ADEC

Item	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	.70	.81	-	.40	.35	.72	.52	.52	.45	.63	.41	.64	.60	.38	.29	.29	.51	.57
2	.71	.83	.46	-	.18	.33	.33	.55	.51	.31	.30	.32	.44	.41	.22	.18	.38	.46
3	.28	.55	.45	.33	-	.31	.22	.28	.11	.36	.32	.23	.10	.09	.03	.18	.19	.44
4	.66	.83	.84	.39	.42	-	.57	.53	.46	.60	.34	.65	.54	.42	.21	.23	.44	.54
5	.52	.71	.62	.46	.40	.66	-	.48	.38	.53	.25	.56	.49	.43	.21	.21	.41	.41
6	.74	.84	.66	.60	.42	.68	.61	-	.60	.45	.38	.54	.53	.54	.18	.25	.56	.49
7	1.1	.93	.57	.67	.28	.60	.59	.74	-	.44	.15	.50	.54	.50	.15	.20	.55	.40
8	.87	.84	.76	.38	.45	.74	.64	.58	.54	-	.38	.57	.53	.33	.20	.21	.41	.46
9	.34	.69	.57	.46	.53	.49	.49	.52	.37	.51	-	.40	.31	.26	.18	.22	.30	.39
10	.71	.88	.77	.41	.34	.79	.70	.69	.65	.70	.58	-	.64	.53	.27	.28	.50	.39
11	.67	.85	.74	.49	.22	.65	.62	.67	.69	.64	.47	.78	-	.47	.28	.23	.57	.40
12	.88	.93	.49	.49	.16	.56	.57	.69	.62	.43	.41	.68	.63	-	.12	.19	.43	.39
13	.44	.73	.41	.33	.23	.29	.45	.27	.33	.25	.41	.42	.43	.20	-	.45	.33	.20
14	.40	.69	.33	.25	.36	.26	.35	.27	.30	.18	.36	.34	.29	.17	.62	-	.29	.24
15	.77	.93	.62	.52	.39	.55	.61	.67	.76	.50	.54	.62	.65	.55	.47	.40	-	.40
16	.64	.76	.63	.57	.70	.62	.58	.59	.55	.52	.61	.49	.48	.49	.38	.32	.55	-

Concurrent validity. As indicated in Table 6, the ADEC was strongly correlated with the ADOS and ADI-R and their subscales (with the exception of the ADI-R Restricted and Repetitive Behaviours score), with rs ranging from .47 to .86.

The concurrent validity of the ADEC was also assessed by comparing how well the ADEC classification agreed with the ADOS classification of AD and DSM-IV-TR diagnosis of AD done at initial assessment. Because the ADEC was designed to screen specifically for AD, rather than PDD-NOS, children classified as PDD-NOS based on the DSM-IV-TR were excluded from this analysis. Cohen's Kappa indicated significant agreement between the ADEC classification and the ADOS classification of AD ($k = .66$, $N = 41$, $p < .001$) and DSM-IV-TR diagnosis of AD ($k = .86$, $N = 191$, $p < .001$) for the unmatched sample. For the matched sample, there was also significant agreement between the ADEC classification and the ADOS classification of AD ($k = .61$, $N = 20$, $p = .003$) and DSM-IV diagnosis of AD ($k = .74$, $N = 78$, $p < .001$).

Table 5

Factor Matrix for PAF of One Factor Solution of ADEC Items

Item	Factor Pattern coefficients	Communalities
	Factor 1 – Social Communication	
1. Response to name	**.80**	.64
2. Gaze monitoring	**.78**	.60
3. Joint attention and social referencing	**.77**	.59
4. Functional play	**.75**	.56
5. Following verbal commands	**.74**	.54
6. Reciprocity of smile	**.70**	.49
7. Use of gestures	**.68**	.46
8. Eye contact	**.66**	.43
9. Pretend play	**.66**	.43
10. Task switching	**.66**	.43
11. Delayed language	**.60**	.36
12. Imitation	**.57**	.32
13. Reaction to common sounds	**.48**	.23
14. Nestling	.36	.13
15. Ritualistic play	.35	.12
16. Anticipation of social advances	.33	.11

Note: Major loadings above .40 are bolded.

Table 6

Concurrent Validity of ADEC

Measure	1	2	3	4	5	6	7	8	9	10	11
1. ADEC	-										
2. ADOS Comm	.76**	-									
3. ADOS RSI	.84**	.77**	-								
4. ADOS Play	.47**	.58**	.48**	-							
5. ADOS SBRI	.60**	.52**	.40**	.24	-						
6. ADOS Rev	.86**	.88**	.95**	.50**	.60**	-					
7. ADOS Severity	.77**	.86**	.89**	.45**	.60**	.96**	-				
8. ADI-R RSI	.58**	.39	.59**	.22	.04	.54**	.54*	-			
9. ADI-R Comm	.63**	.53**	.61**	.53**	.10	.60**	.45*	.43*	-		
10. ADI-R RRB	.16	-.24	-.11	-.38	-.01	-.17	-.17	.02	-.03	-	
11. ADI-R Total	.72**	.42	.64**	.26	.07	.58**	.52*	.88**	.73**	.28	-

Note: ADOS Comm Autism Diagnostic Observation Schedule Communication, *ADOS RSI* Reciprocal Social Interaction, *ADOS SBRI* Stereotyped Behaviours and Restricted Interests, *ADI-R RSI* Autism Diagnostic Interview – Revised Reciprocal Social Interaction, *ADI-R Comm* Communication, *ADI-R RRB* Restricted and Repetitive Behaviours
** $p < .01$ (2-tailed)
* $p < .05$ (2-tailed)

Diagnostic validity. A one-way analysis of variance (ANOVA) compared the different diagnostic groups, based on the DSM-IV-TR diagnoses (i.e., AD, PDD-NOS and non-ASD), on the ADEC total score to determine diagnostic validity. As the group difference in NVIQ approached significance, F (2, 65) = 2.69, p = .08, an analysis of covariance (ANCOVA) was performed with the NVIQ specified as a covariate. A preliminary analysis evaluating the homogeneity-of-slopes assumption indicated that the relationship between the NVIQ and the ADEC score did not differ significantly as a function of the group factor, $F(2, 62) = 3.03$, $p = .06$, $\eta2 = .09$. The ANCOVA did, however, indicate significant group differences for mean ADEC total scores, $F(2, 64) = 24.17$, $p <$ .001, $\eta2 = .43$, after controlling for NVIQ. Post hoc comparisons using the Games-Howell procedure (due to unequal variances across the groups) showed significant differences between the groups of AD and PDD-NOS ($p < .001$) and AD and non-ASD ($p < .001$), with participants in the AD group scoring the highest, followed by participants with PDD-NOS and the non-ASD group the lowest.

Diagnostic validity of the ADEC was also examined for children below age 24 months as other screening tools such as the CARS and STATS are considered unsuitable for children below this age group. The sample included 14 children with AD, 6 with PDD-NOS, 30 with non-ASD and 36 TD children aged between 12 to 24 months old. As there was no significant group difference in the NVIQ ($p = .65$), an ANOVA on the ADEC scores was conducted for this sample and results indicated a significant group difference, $F(3, 82) = 33.99$, $p <$ .001, $\eta2 = .55$, with participants in the AD group scoring higher than the non-ASD group ($M = 18.0$, $SD = 5.5$ vs. $M = 8.7$, $SD = 5.8$).

Next, ADEC scores for the AD versus the non-ASD group were examined, but with the Vineland Adaptive Behavior Composite (VABC) specified as a covariate instead of the NVIQ. Young children with suspected ASD often present an assessment challenge due to their social difficulties, unusual use of language, frequent off-task behaviours, high distractibility and variable motivation (Chawarska & Bearss, 2008; Ozonoff, Goodlin-Jones, & Solomon, 2005) and their relatively low scores on the MSEL (and Bayley III) may not be a true reflection of their abilities. The relationship between the VABC and the ADEC score did not differ significantly as a function of the group factor, $F(1, 91) = .13$, $p = .72$, $\eta2 = .00$. Moreover, the ANCOVA indicated significant group ADEC differences, $F(1, 92) = 35.43$, $p < .001$, $\eta2 = .28$, with participants in the AD group scoring higher than the non-ASD group.

Finally, an ANOVA on ADEC scores for the matched AD and non-ASD samples (described earlier) also indicated a significant group difference, $F(1, 76) = 48.57$, $p < .001$, $\eta2 = .39$, with participants in the AD group scoring higher than the non-ASD group.

ROC Analysis

The screening properties of the ADEC were examined using the signal detection procedure of ROC curve analyses. This analysis was used to identify the proportion of children with AD who are correctly identified as at risk (i.e., sensitivity), and the proportion of children without AD who are correctly identified as *not* at risk (i.e., specificity), associated with the different cutoff scores. In addition, whether the cutoff score of 11 (recommended in the ADEC manual) was the optimal score was also examined in this study. Finally the area

under the curve (AUC - a measure of the overall predictive validity) and whether it exceeded .90, indicating excellent validity was examined.

Signal detection analyses were conducted separately for (1) unmatched samples of children in the (a) AD *versus* non-ASD groups and (b) AD *versus* non-ASD combined with TD groups, and (2) matched samples of children in the (a) AD *versus* non-ASD groups and (b) AD *versus* non-ASD combined with TD groups. It has been suggested that specificities usually improve when TD cases are combined with the non-ASD cases (Kim & Lord, 2012a) and I examined whether this was the case. As illustrated in Table 7, in all 4 cases the optimal cutoff score was 11, sensitivity was 1.0 while specificity ranged from .74 to .90 and the AUC ranged from .87 to .95.

Positive and negative predictive values (PPV and NPV) were also calculated. PPV measures the proportion of children who screen positive (i.e., at risk) who actually have AD and NPV measures the proportion of children who screen negative (i.e., not at risk) who do not have AD. In the unmatched sample with a cutoff score of 11, PPV was .84 (i.e., 70/83 true positives) and NPV was 1.0 (i.e., 108/108 true negatives). Thus, 13 children were over-identified as having AD (false positive) and no children with AD were missed (false negative). One of the false positive cases obtained the minimum cut-off score of 11 on the ADEC, though clinical judgment deemed that he did not meet diagnostic criteria for any ASD. Ten false positive cases were diagnosed with severe developmental delay (non-verbal IQ below 35). The last two false positive cases were a 12-month old child with an ADEC score of 15 and a 15-

Table 7

Sensitivity and Specificity Associated with the Different ADEC Cutoff Score Based on Different Samples

ADEC Initial Score	Unmatched AD (*n*=70) versus Non-ASD (*n*=57)			Unmatched AD (*n*=70) versus Non-ASD+TD (*n*=121)			Matched AD (*n*=39) versus Non-ASD (*n*=39)			Matched AD (*n*=39) versus Non-ASD+TD (*n*=103)		
	Sens	Spec	AUC (95% CI)	Sens	Spec	AUC (95% CI)	Sens	Spec	AUC (95% CI)	Sens	Spec	AUC (95% CI)
3	1.0	.11	-	1.0	.37	-	1.0	.15	-	1.0	.52	-
5	1.0	.28	-	1.0	.53	-	1.0	.31	-	1.0	.66	-
7	1.0	.42	-	1.0	.66	-	1.0	.41	-	1.0	.73	-
9	1.0	.61	-	1.0	.78	-	1.0	.59	-	1.0	.81	-
10	1.0	.68	-	1.0	.82	-	1.0	.62	-	1.0	.85	-
11	**1.0**	**.77**	**.90 (.84, .96)**	**1.0**	**.89**	**.95 (.92, 9.8)**	**1.0**	**.74**	**.87 (.79, .95)**	**1.0**	**.90**	**.87 (.79, .95)**
12	.93	.79	-	.93	.90	-	.90	.74	-	.90	.90	-
13	.86	.81	-	.86	.91	-	.85	.74	-	.85	.90	-
15	.74	.83	-	.74	.92	-	.72	.77	-	.72	.80	-

*Note: Sens=*Sensitivity, *Spec=*Specificity *AUC=*Area Under the Curve

month old child with an ADEC score of 12. In sum, the ADEC accurately

identified the low-functioning children with AD, but was over-inclusive. When

the ten cases with severe delay were removed, the specificity improved from .89

to .96 and the PPV also improved from .84 to .95.

Given the large variability in non-verbal IQ, ROC analyses were also

used to examine whether the ADEC was any better at differentiating AD from

non-ASD than developmental or adaptive behaviour assessments such as the

Mullen's or the Vineland. In this case, when the unmatched sample of AD

versus combined TD and non-ASD cases was examined, though the optimal

sensitivity was 1.0, the specificity was lower at .07 using a corresponding NVIQ

score of 35. The AUC was only .12, 95% CI [.01, .22] which indicated that the

test does worse than chance at correctly identifying children with AD. When the

VABC was used instead of the NVIQ, the AUC was even lower at .08, 95% CI

[.00, .17], with the optimal sensitivity of .92 and sensitivity of .03 corresponding

to the VABC score of 54. Thus, although all of these tests were able to detect

atypical development, they were unable to identify autism per se. Overall the

data in Study 1 showed that the ADEC was detecting more than disability or low

NVIQ and/or VABC. Specifically, it was able to identify the children with AD

and without AD.

Discussion

This study demonstrated that the ADEC is a psychometrically sound and

effective Level 2 AD screening tool suitable for use with young children ranging

from 12 to 36 months of age. The ADEC scores were reliable across examiners

and across testing occasions, and internal consistency was high. The findings

also demonstrate the construct, concurrent and diagnostic validity of the ADEC.

The ADEC showed high concurrent validity with the ADOS, ADI-R and DSM-IV-TR classification of AD, and diagnostic validity across the different groups (AD, PDD-NOS and non-ASD), even after controlling for participants' NVIQ and Vineland Adaptive Behavior Composite scores. Diagnostic validity of the ADEC for a sub-group of children aged 24 months and below was also demonstrated. In addition, the ADEC provided high sensitivity (1.0), specificity (.89) and predictive values (PPV = .84, NPV = 1.0) using a cutoff score of 11 in identifying young children referred for possible risk for AD with the validation sample, and was clearly superior in these areas to an IQ or adaptive functioning test. This value of 11 is also consistent with the ADEC manual's interpretation of 'moderate risk of autism' where the objective is to maximise the chance of a child with possible AD being correctly identified while, at the same time, minimising the risk of over-including those with other forms of developmental delay.

Even though the ADEC was not designed as a Level 1 screening tool (i.e., population screening), I have included some typically developing children in the data analysis as research with younger children often contrasts AD and mixed TD and non-ASD samples (as in studies with baby siblings) (Kim & Lord, 2012a). When the TD and non-ASD cases were combined in this study's analysis, it was found that specificity was improved when compared to just using the non-ASD cases alone. Specificity was slightly lower (.89) when individuals with severe levels of intellectual disability were included; removing those cases improved specificity to .96 and PPV also improved from .84 to .95. Other studies on diagnostic and screening instruments for ASD also over-identified individuals with severe intellectual disability (e.g., De Bildt et al.,

2004; Gray et al., 2008; Maljaars, Noens, Scholte, & van Berckelaer-Onnes, 2012; Witwer & Lecavalier, 2007) because some ASD behaviours such as absence of language development and imagination are also observed in individuals with severe or profound intellectual disability (Matson & Shoemaker, 2009; Wilkins & Matson, 2009). When compared to the other Level 2 screening tools (e.g., the CARS and STAT), the ADEC performed at a similar level (in terms of sensitivity and specificity) to them. In this case, the specificity of the ADEC was higher when individuals with severe levels of intellectual disability were excluded.

A factor analytic examination of the construct validity of the ADEC indicated a one-factor solution (Social Communication) which showed excellent internal consistency. Although this one-factor solution of autistic symptoms contrasts with other two or three factor models (e.g., Frazier, Youngstrom, Kubu, Sinclair, & Rezai, 2008; Snow, Lecavalier, & Houts, 2009), it is not the first time this has been proposed (Constantino et al., 2004). Further, the extraction of a factor combining social and communication criteria is consistent with the new DSM criteria that propose combining these domains (*DSM-5;* APA, 2013). In addition, these data cause doubt about the value of the DSM-IV-TR third domain (restricted, repetitive and stereotyped behaviour) for screening in young children. It has previously been argued that these behaviours may not emerge until later in development (Charman & Baird, 2002; Eaves & Ho, 2004; Moore & Goodson, 2003) and thus it may be these behaviours should not be considered when assessing very young children. Investigation of the prevalence of these behaviours in very young children is thus warranted.

Impaired social interaction is a defining feature of ASD for all age ranges. Indeed, across ASD subtypes, impairment in the social interaction domain (whether accompanied by impairment in communication and/or restricted, repetitive and stereotyped behaviour) criterion must be fulfilled in order for an individual to be classified as ASD (APA, 2000). Screening efforts for very young children should attend to very basic social interaction abilities with emphasis on the earliest signs of impairment that seem to distinguish ASD from general developmental delay (i.e., joint attention, response to name, eye contact). Thus, those ADEC items identified in the Social Communication factor might be able to be used to detect these core social deficit behaviours in toddlers with ASD.

Clinical Implications

This validation study adds to the currently limited literature of using the ADOS revised algorithm scores (Gotham et al., 2007) and the new ADOS severity score (Gotham et al., 2009) in relation to an autism-specific screening instrument, in this case the ADEC. There was a significant and strong correlation between the ADEC scores and the ADOS revised algorithm total scores ($r = .86$) and the ADOS severity score ($r = .77$). In addition, there was also a significant agreement between the ADEC classification and the ADOS classification of AD. These data indicate that the ADEC can be used with confidence to identify children at risk of AD. Nevertheless, it should be noted that the ADEC is designed for young children under 3 years of age and may not be suitable or effective in screening older children and adolescents suspected of having AD.

Besides using the ADEC as a Level 2 screening tool for AD, the ADEC can also provide important prescriptive information for early intervention programming. For example, studies have indicated the importance of motor imitation skills, initiating joint attention, play skills and responding to joint attention at early age which predicted better language outcomes at later age (e.g., Kasari, Gulsrud, Freeman, Paparella, & Hellemann, 2012; Stone & Yoder, 2001; Thurm, Lord, Lee, & Newschaffer, 2007). These are some of the core deficit behaviours that are easily identified in the ADEC that can be used for planning individualized educational programs for a child with ASD.

Limitations

There are some limitations of this study that should be considered. First, all the children with AD in this sample were found to have intellectual disability (based on IQ < 70). Hence, it is unclear how the children with AD, but no developmental delay, will perform on the ADEC. It is worth noting, however, that about one-fifth (17.5%) of the non-ASD group had *severe* intellectual disability, which resulted in the over-inclusion of these children as contrasted to the usual 3 to 4 percent of the intellectually disabled population which has *severe* intellectual disability (APA, 2000) and hence may have affected the sensitivity and specificity of the ADEC. For example, it was found in this study that specificity was improved from .89 to .96 when these individuals with severe levels of intellectual disability were removed from the ROC analysis. Hence replication with larger samples (ideally with more of the participants with mild and moderate intellectual disability which is representative of the distribution in the intellectually disabled population) should be carried out to evaluate the generalizability of the present findings.

Second, even though the current "gold standard" diagnostic tools (ADOS and ADI-R) were used, they are not designed for younger children (e.g., below 24 months). For example, the ADOS tends to over-classify children with other developmental difficulties as having AD or ASD when clinical judgment deems that they do not (Gotham et al., 2007). In addition, not all participants (especially the non-ASD) were administered the ADI-R. To resolve this issue, the decision of making BEC diagnosis was chosen because it uses all sources of information rather than the ADOS or ADI-R alone. Nevertheless, with the availability of the ADOS Toddler module (Luyster et al., 2009) and the ADI-R Toddler (Kim & Lord, 2012b), future research can be undertaken to compare the ADEC results against these instruments that are specifically designed for younger children[3].

Lastly, demographic information (e.g., SES and educational level) was not available for most of the participants. However, it was not the aim in this study to speculate about family variables that might influence parents' awareness of ASD and whether that prompted their willingness to participate in my screening study.

It is important to realise that the ADEC was designed to identify children specifically with AD, rather than for all forms of ASD. At the time of data collection and writing, the new DSM-5 criteria have not been released yet. Hence, in this validation study, children with PDD-NOS were excluded from analyses that examined the rate of agreement between the ADEC and ADOS and DSM-IV-TR and ROC curves. Consequently, it may not be feasible to use the

---

[3] Further result about the ADOS Toddler and the ADI-R Toddler in relation to the ADEC will be discussed in Study 3.

proposed ADEC cut-off score for AD to identify these children with PDD-NOS or the higher functioning children.

Though one of the objectives of this validation study was to examine the ADEC against the DSM-IV-TR criteria, it was also critical to explore the ADEC in the light of the new DSM-5 changes (APA, 2013). Hence the next two studies investigated the use of the ADEC on children classified with a DSM-5 ASD diagnosis; specifically Study 2 examined how well the ADEC predicted long term outcomes such as DSM-IV-TR and DSM-5 ASD classifications in children while I developed a brief ADEC version to identify children with a DSM-5 ASD diagnosis in Study 3.

CHAPTER 3

Study 2[4]: Using the Autism Detection in Early Childhood (ADEC)

and Childhood Autism Rating Scales (CARS) to Predict Long Term

Outcomes in Children with Autism Spectrum Disorders

There has been considerable clinical and research interest in developing

appropriate screening instruments to help identify young children at risk of

developing an ASD, particularly as there is currently no biological marker to

assist with identification (Abrahams & Geschwind, 2008). Although most of the

widely used ASD Level 2 screening tools have provided sufficient data about

some of their psychometric properties (e.g., test-retest reliability, internal

consistency and concurrent validity; e.g., Matson et al., 2007; Matson et al.,

2011; Stone et al., 2000; Stone et al., 2004), there is little information available

about the predictive validity of these tools, despite this being an integral part of

the validation process of a screening tool.

Establishing the predictive validity of ASD screening tools is important,

as it allows clinicians to use information obtained in the administration of the

screening test to determine (a) how likely it is that these children will continue

to meet diagnostic criteria for ASD with age and, more importantly, (b) the

likely severity of this disorder. It is known that ASD can be reliably identified

and diagnosed in young children as young as 24 months of age (Moore &

Goodson, 2003) and the diagnosis may remain quite stable. For example, Lord

---

[4] Material presented as part of this study appeared in the *Journal of Autism and Developmental Disorders* as 'Using the Autism Detection in Early Childhood (ADEC) and Childhood Autism Rating Scales (CARS) to Predict Long Term Outcomes in Children with Autism Spectrum Disorders' (Nah et al., 2014*).*

et al.'s (2006) longitudinal study using standardised diagnostic tools such as the

Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) and the

Autism Diagnostic Interview - Revised (ADI-R; Le Couteur et al., 2003) found

that diagnosis of AD in 2-year-olds was quite stable up to 9 years of age.

Study 2 examines the predictive validity of two of the commonly used

Level 2 ASD screening tools: the Childhood Autism Rating Scale (CARS;

Schopler et al., 1988) and the Autism Detection in Early Childhood (ADEC;

Young, 2007). There is only one other study that has examined the predictive

validity of a screening instrument, the Parent Observation of Early Markers

Scale (POEMS; Feldman et al., 2012). Preliminary results were promising with

respect to diagnostic classifications obtained one to two years after initial

assessment. The POEMS was, however, designed as a parent report measure to

monitor the behavioural development of infants at risk for ASD and there

appears to be no clinician-administered screening tool with predictive validity

established for long term outcomes reported.

Currently, there is little information available about the validity of Level

2 screening tools in predicting long term outcomes such as diagnostic

classification. The ADEC and the CARS were examined in this study as both

are clinician-administered ASD screening tools suitable for use with young

children. Detailed psychometric evaluation of the ADEC was provided in Study

1 although no data on the long term predictive validity of the ADEC was

available at the time of that study. The CARS is arguably the most established

and widely used Level 2 screening tool for the identification of ASD (Luiselli et

al., 2001; Ozonoff et al., 2005) and is suitable for use with children over two

years of age. The CARS has been compared with other instruments such as the

Autism Diagnostic Interview - Revised (ADI-R; Pilowsky, Yirmiya, Shulman,
& Dover, 1998), the Autism Treatment Evaluation Checklist (ATEC; Geier,
Kern, & Giier, 2013) and the Autism Behavior Checklist (ABC; Rellini,
Tortolani, Trillo, Carbone, & Montecchi, 2004) in terms of diagnostic validity.
There have been no such comparisons involving the ADEC, nor have there been
any studies evaluating the predictive validity of either the CARS or the ADEC
for long term outcomes.

Study Aims

This study examines the predictive validity of two Level 2 ASD
screening tools (the ADEC and CARS) in relation to diagnostic classifications,
symptom severity and functioning level at 2 and 6 years following initial
assessment. To date, no studies have provided this comprehensive evaluation
(and comparison) of the predictive validity of Level 2 ASD screening
instruments on long term outcome measures in children with ASD.

Method

Participants

The participants in this study were all involved in an intensive
behavioural intervention program using a structured applied behavioural
analysis (ABA) framework (see Young, Partington, & Goren, 2009) at a
university autism research center in South Australia. Data collection was spread
over a 10-year period from 2003 to 2013, and the author was involved in the
(six-year follow up assessment) data collection[5] for this study from April 2011
to January 2012. The program is a free program open to any participants. The

---

[5] My role includes designing the study, contacting participants, administering various tests such
as the ADOS and WISC-IV to the participants, data analysis and write up.

ethnic background of the sample was predominantly Caucasian (96.4 %). Only

15 participants' data (27.2 %) about SES and parental education level were

available. Demographic information was not available for the earlier batch of

participants who first joined the research intervention program and some

participants chose not to provide the details. Of these 15 families, the primary

care-giver was the mother and they were generally well educated, with 73.3 %

having at least some diploma education. The average total family income fell in

the $AUD 60,000–$100,000 range.

Participants' families were made aware of the program at either at

diagnosis or upon registration with the local Autism Association. Participants

completed a two week on campus intervention program and their parents/care-

givers were also trained to carry out the home based intervention program with

their child in the next 18 weeks. During the 18 week period, the parents would

meet a staff member from the intervention program fortnightly for a follow-up

so that the staff could look over the program, give feedback and adjust the

program as necessary, as well as address any questions or concerns. The parents

were encouraged to continue the behavioural intervention program on their own

after the 18 week period. All contactable participants were invited to return for 2

year and 6 year follow-up assessment. There were 55 children (44 male, 11

female) aged between 19 to 42 months old ($M = 33.5$, $SD = 5.6$) available for

the initial assessment in this study. Within this sample, 27 (49.1%) children

were verbal while 28 (50.9%) children were non-verbal (i.e., not using any

words at all).

Thirty-seven of the 55 participants in this study were part of the

participant sample in Study 1 while the remaining 18 participants in the present

sample were not eligible for inclusion in Study 1's psychometric evaluation of

the ADEC because they were older at initial assessment than the age for which

the ADEC was intended. An ANOVA was conducted to examine whether there

was any difference in ADEC and CARS scores between the 37 participants

(aged below 36 months) and the 18 participants (aged above 36 months).

Results indicated no significant differences in either ADEC scores ($M = 17.0$,

$SD = 5.8$ vs. $M = 14.6$, $SD = 6.6$), $F(1, 53) = 1.93$, $p = .17$, $\eta2 = .04$, or CARS

scores ($M = 34.4$, $SD = 5.0$ vs. $M = 34.1$, $SD = 3.0$, $F(1, 53) = 0.05$, $p = .82$, $\eta2$

$= .00$, between these two group of participants.

A best estimate clinical (BEC) DSM-IV-TR diagnosis was made by the

author using all available information and assessment results (including

behavioural descriptions that were reported by parents in the CARS interview),

excluding ADEC data and CARS individual item score and total score, to

generate diagnoses independent of the ADEC. Of the 55 children, 51 received

an initial BEC DSM-IV-TR clinical diagnosis of Autistic Disorder (AD), 2

received a diagnosis of Pervasive Developmental Disorder-Not Otherwise

Specified (PDD-NOS) and 2 a diagnosis of non-autism spectrum disorders (non-

ASD) (such as speech and developmental delay) according to DSM-IV-TR

criteria. Within this sample, 78.2% had an independent confirmatory diagnosis

(i.e., a separate diagnosis apart from the researcher's BEC diagnosis) from either

two other independent professionals who were recognized by the state's autism

association or other professionals such as paediatricians and psychologists. The

author was blind to the results of this confirmatory diagnosis. The inter-rater

reliability for DSM-IV-TR diagnosis of AD between the researcher's BEC

diagnosis and the independent diagnosis was substantial ($k = .64$, $p < .001$) based on Landis and Koch (1977)'s guidelines.

Fifty-three of the 55 children (96.4%) were available for the two-year follow up assessment and 22 children (40%) were located and followed up after about 6 years ($M = 68.6$ months, $SD = 17.9$). Two participants withdrew from the two-year follow up assessment; there were no significant differences between participants who were and were not involved in the two-year follow up assessment in terms of gender, $\chi^2 (1) = .52$, $p = .47$, initial age, $F(1, 53) = 1.40$, $p = .24$, $\eta2 = .03$, or initial scores on CARS, $F(1, 53) = .28$, $p = .87$, $\eta2 = .00$ or ADEC, $F(1, 53) = .03$, $p = .60$, $\eta2 = .01$. At the six-year follow up assessment, 10 participants did not wish to participate, 14 participants were not contactable or moved interstate, and 9 participants were not yet eligible for the 6 year follow up. There were no significant differences between participants who participated and did not participate in the six-year follow up assessment in terms of gender, $\chi^2 (1) = .93$, $p = .34$, initial age, $F(1, 53) = 1.08$, $p = .30$, $\eta2 = .04$, initial ADEC score $F(1, 53) = 3.81$, $p = .06$, $\eta2 = .04$ and initial CARS score $F(1, 53) = 1.06$, $p = .31$, $\eta2 = .04$. Sample characteristics are presented in Table 8.

Intake Measures

CARS. The CARS consists of 14 domains assessing behaviours associated with autism such as adaptation to change, listening response, verbal communication, and relationship to people, as well as a 15th domain rating general impressions of autism. The child's behaviour is rated on a scale based on deviation from the typical behaviour of children of the same age. Each domain is scored on a scale ranging from one to four, with higher scores associated with a higher level of impairment. The CARS has good internal consistency ($\alpha = .91$-

.94) (Schopler et al., 1988; Tachimori, Osada, & Kurita, 2003), an intraclass

correlation for the Total score of .83 to .94 (DiLalla & Rogers, 1994; Perry &

Freeman, 1996) and test-retest stability over a one-year period of .88 for the

Total score (Schopler et al.). Eaves and Milner (1993) reported a sensitivity of

.98 when using the CARS cut-off score of 30 for diagnoses of AD.

Table 8

Sample characteristics

	Initial assessment	Two-year follow up	Six-year follow up	N
N (male, female)	55 (44, 11)	53 (42, 11)	22 (19, 3)	-
Age (months)	33.5 (5.6)	58.6 (7.2)	103.8 (19.6)	-
CARS score	34.3 (4.4)	-	-	55
ADEC score	16.2 (6.1)	-	-	55
DSM-IV-TR AD	51 (92.7%)	44 (83%)	14 (63.7%)	-
DSM-IV-TR PDD-NOS	2 (3.6%)	7 (13%)	5 (22.7%)	-
DSM-IV-TR non-ASD	2 (3.6%)	2 (4%)	3 (13.6%)	-
DSM-5 ASD	-	45 (84.9%)	16 (72.7%)	-
DSM-5 non-ASD	-	8 (15.1%)	6 (27.3%)	-
Vineland ABC	-	58.6 (14.5)	-	46
ABAS-II GAC	-	-	62.1 (19.2)	18
ADOS Revised	-	-	14.1 (7.7)	22
ADOS Severity	-	-	6.1 (2.8)	22
SCQ	-	-	11.7 (5.6)	21
WISC-IV PRI	-	-	80.1 (18.7)	17

Note: CARS Childhood Autism Rating Scale, *ADEC* Autism Detection in Early Childhood, *AD* Autistic Disorder, *PDD-NOS* pervasive developmental disorder - not otherwise specified, *NS* non-autism spectrum disorders, *ASD* Autism Spectrum Disorder, *Vineland ABC* Vineland Adaptive Behavior Composite standard score, *ABAS-II GAC* Adaptive Behavior Assessment System – 2nd Ed General Adaptive Composite standard score, *ADOS* Autism Diagnostic Observation Schedule, *SCQ* Social Communication Questionnaire Total Score, *WISC-IV PRI* Wechsler Intelligence Scale for Children 4th Ed Perceptual Reasoning Index

ADEC. The ADEC has been found to be a reliable and valid screening tool (Study 1). To summarise the findings in Study 1, I found concurrent validity of the ADEC with other more intensive instruments such as the ADOS (Lord et al., 2000), the ADI-R (Le Couteur et al., 2003) and clinical judgment based on DSM-IV-TR diagnosis (APA, 2000). The ADEC also has high sensitivity (1.0), specificity (.89) and predictive values (PPV = .84, NPV = 1.0) when using a cutoff score of 11 in identifying young children referred for possible risk for AD with the validation sample.

Outcome Measures at 2-Year Follow Up Assessment

Overall Adaptive Functioning. The Vineland Adaptive Behavior Scales (Vineland et al., 1984) was used as a measure of overall adaptive functioning level. Details about the Vineland were provided in Study 1. The Adaptive Behavior Composite (ABC) was used as overall adaptive outcome.

Diagnostic Classification. At the time of writing, the ADEC has been validated for children only with a DSM-IV-TR AD diagnosis and not with the broader autism spectrum. Therefore, I examined diagnostic classification based on the previous DSM-IV-TR (strictly AD) and the new DSM-5 ASD (which includes both the AD and other forms of ASDs) (APA, 2013). A best estimate clinical (BEC) DSM-IV-TR and the DSM-5 ASD classifications diagnosis was

made by the author using all available information and assessment results, excluding ADEC data, to generate diagnoses independent of the ADEC. Inter-rater reliability was taken for a subset of this sample (20%) by an experienced clinician in ASD assessment. The inter-rater reliability for DSM-5 diagnosis of ASD between the researcher's BEC diagnosis and the independent diagnosis was perfect ($k = 1.0$, $p < .005$).

Outcome Measures at 6-Year Follow Up Assessment

      Overall Adaptive Functioning. The Adaptive Behavior Assessment System – 2nd Ed. (ABAS-II; Harrison & Oakland, 2003) was used as a measure of overall adaptive functioning level. Similar to the Vineland, the ABAS-II is a standardised, individually administered assessment of strengths and weaknesses in adaptive functioning in persons aged from birth to adulthood. Because the ABAS-II is the only instrument that provides standardised scores according to both the 10 adaptive skills areas defined by the DSM-IV-TR: communication, community use, functional academics, home-living, health and safety, leisure, self-care, self-direction, social, and work; and the three adaptive behaviour domains (conceptual, practical, and social skills) defined in the 11th edition of the AAIDD terminology and classification manual (Schalock et al., 2010), I used the ABAS-II here. However, it should be noted that the aim of this study was to compare the initial ADEC score to the adaptive score at the six-year follow up assessment rather than comparing the adaptive score at the two time frames. The ABAS-II also provides an index of overall adaptive behaviour/developmental maturity- General Adaptive Composite (GAC). Estimates of internal consistency and test-retest reliability for the GAC are in the .90s (Harrison & Oakland). Correlations between the Adaptive Behavior

Composite on the Vineland Adaptive Behavior Scale and the ABAS-II GAC

ranged from .70 to .84 (Rust & Wallace, 2004).

ASD Symptomatology. The ADOS (Lord et al., 2000) was chosen as a

measure of ASD symptomatology as it is widely recognized as a gold standard

assessment of communication, social interaction, and play or imaginative use of

materials for individuals suspected of having AD or ASD. Besides using the

revised ADOS algorithm total scores as a stand-in for a measure of autism

severity (Gotham et al., 2007), the total scores from the revised ADOS

algorithms were also used to provide a continuous measure of overall ASD

symptom severity (on a 10-point scale) that is less influenced by child

characteristics, such as age and language skills, than raw totals (Gotham et al.,

2009). Lower severity scores are associated with less autism impairment.

The Social Communication Questionnaire (SCQ; Rutter et al., 2003) was

another measure of ASD symptomatology used in this study. The SCQ evaluates

communication skills and social functioning in children over 4 years of age with

an equivalent mental age of at least 2 years who may have ASD. The SCQ is

based on the ADI-R (Le Couteur et al., 2003) and is completed by a parent or

other primary caregiver in about 10 minutes, and focuses on the child's

developmental history and their behaviour over the last 3 months. Total scores

can range from 0 to 39. Four levels are reported: low (< 8); moderately low (8-

14); moderately high (15-21); high (> 22). The cutoff for ASD is ≥ 15, and a

cut-off for autism is also reported (22). Internal consistency reliability of the

SCQ was .90 (Berument, Rutter, Lord, Pickles, & Bailey, 1999).

Non-verbal Cognitive Functioning. The Wechsler Intelligence Scale

for Children (WISC-IV; Wechsler, 2003) is a clinically administered

intelligence (IQ) test for children aged from 6:0 to 16:11 years. As some children obtained significant discrepancies between the different composite scores or their scores on the Verbal composite score were invalid (due to too many raw score of 0), the Perceptual Reasoning Index (PRI) which is a measure of non-verbal reasoning skill was used instead of the composite Full Scale IQ (FSIQ) score. In addition, children with ASD often present significant discrepancies between their verbal and non-verbal IQ (Joseph et al., 2002) and, thus, the FSIQ score may also not be meaningful. The internal consistency for the WISC-IV PRI composite scale is .92 (Wechsler, 2003).

Diagnostic Classification. A best estimate clinical (BEC) DSM-IV-TR and the new DSM-5 (APA, 2013) ASD classifications diagnosis was made by the author using all available information and assessment results, excluding ADEC data, to generate diagnoses independent of the ADEC.

Procedure

All parents provided informed consent and appropriate ethics approvals were obtained from the University ethics committee prior to conducting this study. During the initial assessment, a clinician within the intervention program assessed the participants using the ADEC while the CARS was completed by either the same or another clinician. In view of the context of being an early intervention program rather than an assessment centre, the participants were not administered a full battery of assessment tools such as the ADOS. In addition, the participants were too young to be administered the WISC-IV at the initial point of intake. At the end of the assessment session, the DSM-IV-TR checklist for AD was also completed by the either the same clinician or both clinicians together. The CARS and ADEC were used as part of an intake assessment in the

context of an early intervention program and thus there was no control condition. Parents were also given the Vineland survey form to fill in and return to the assessor. These tests were repeated after 2 years either by the same or another qualified staff member in the program. Given staff turn over the clinicians at the initial assessment were not the same as the clinicians testing at the two-year follow up assessment. Further, no clinician was directly involved in the therapy. Where possible the clinician remained naïve to the progress of the child and what assessment was being undertaken (i.e., 2 week, 18 week). This was, however, difficult to orchestrate as parents would typically discuss the intervention despite being asked not to.

At the initial and two-year follow up assessment, the CARS and ADEC were administered to the participants at random by one of four different examiners who had university degrees in psychology or psychology-related courses and had received prior training on administration and scoring, though none was qualified to diagnose ASD. In addition to child observations of social interaction and play, detailed developmental/medical histories were obtained. Reports from the participants' medical or/and psychological reports were also reviewed for additional history and information. All this archival information was used to guide decision-making on the BEC DSM-5 diagnosis for the participants.

For the six-year post initial assessments, children were contacted and assessed with another battery of tests at either the university autism research center or the child's home. Not all participants were available for the six-year follow up assessment, with some unable to be contacted and others declining to participate in the follow up studies. The remaining available participants were

administered the ADOS and the WISC-IV while their parents were given the

ABAS-II Parent form and the SCQ Current form to complete for the assessor.

The ABAS-II, ADOS, SCQ and WISC were administered by two examiners

(one of them was the author) who had a Master's degree in psychology and

received training in the administration and interpretation of these tests. These

two examiners were not involved in the administration of the CARS and ADEC

at the initial assessment. There were 22 children (19 male, 3 female), now aged

between 69 and 146 months old ($M = 103.8$, $SD = 19.6$), who completed the

ADOS assessment ($N = 6$ for Module 1, $N = 9$ for Module 2 and $N = 7$ for

Module 3). The ADOS examiner was blind to the results of the previous

assessments and fully trained to the standards of research reliability set by the

developer. Similar to the two-year assessment, detailed

developmental/psychological/medical assessment reports were obtained and

reviewed to guide the DSM-5 diagnosis for the participants.

Results

Diagnostic Consistency after 2 and 6 Years

First, the agreement between ADEC and CARS classifications of AD

conducted at initial assessment and the DSM-IV-TR classification of AD made

at two- and six-year post initial assessments were examined and compared using

Kappa analysis. Because the ADEC and CARS were meant to identify only AD

and not PDD-NOS, only those children who obtained an initial classification of

AD based on the DSM-IV-TR were used in this analysis. Results indicated

statistically significant though poor-fair agreement between the ADEC

classification at initial, and DSM-IV-TR diagnoses of AD ($k = .37$, $N = 51$, $p <$

$.01$) at two-year and DSM-IV-TR AD ($k = .40$, $N = 21$, $p < .05$) at six-year

follow up assessment. For the CARS there was significant and moderate

agreement between the classification at initial and DSM-IV-TR diagnosis of AD

($k = .41$, $N = 51$, $p < .005$) at the two-year but not at the six-year follow up

assessment for DSM-IV-TR AD ($k = .22$, $N = 21$, $p = .43$). The change in

diagnostic status of participants across the three time periods is displayed in

Figure 1.

Relationship between ADEC and CARS, and Symptom Severity and

Functioning Level

Next, the relationship between the ADEC and CARS total scores at

initial assessment and (a) overall adaptive functioning level using the Vineland-

ABC taken at two-year and the ABAS-II GAC taken at six years, (b) ASD

symptomatology based on ADOS revised algorithm total score (ADOS Rev),

ADOS severity score and SCQ, and (c) non-verbal cognitive functioning using

the WISC-IV PRI taken at six-year follow up assessments were examined. As

expected, the ADEC initial score was significantly correlated with all the

outcome measures with the exception of the ADOS severity score ($p = .055$) and

WISC-IV PRI score ($p = .06$). There was no significant correlation between the

CARS initial score and the other outcome measures except for the Vineland

ABC score. However, there was overlap between the confidence interval (CI) of

the correlation coefficients for the ADEC, 95% CI [.07, .75] and the CARS,

95% CI [-.05, .69] in relation to the ADOS revised algorithm total score, and the

ADEC, 95% CI [.02, .73] and the CARS, 95% CI [-.11, .65] in relation to the

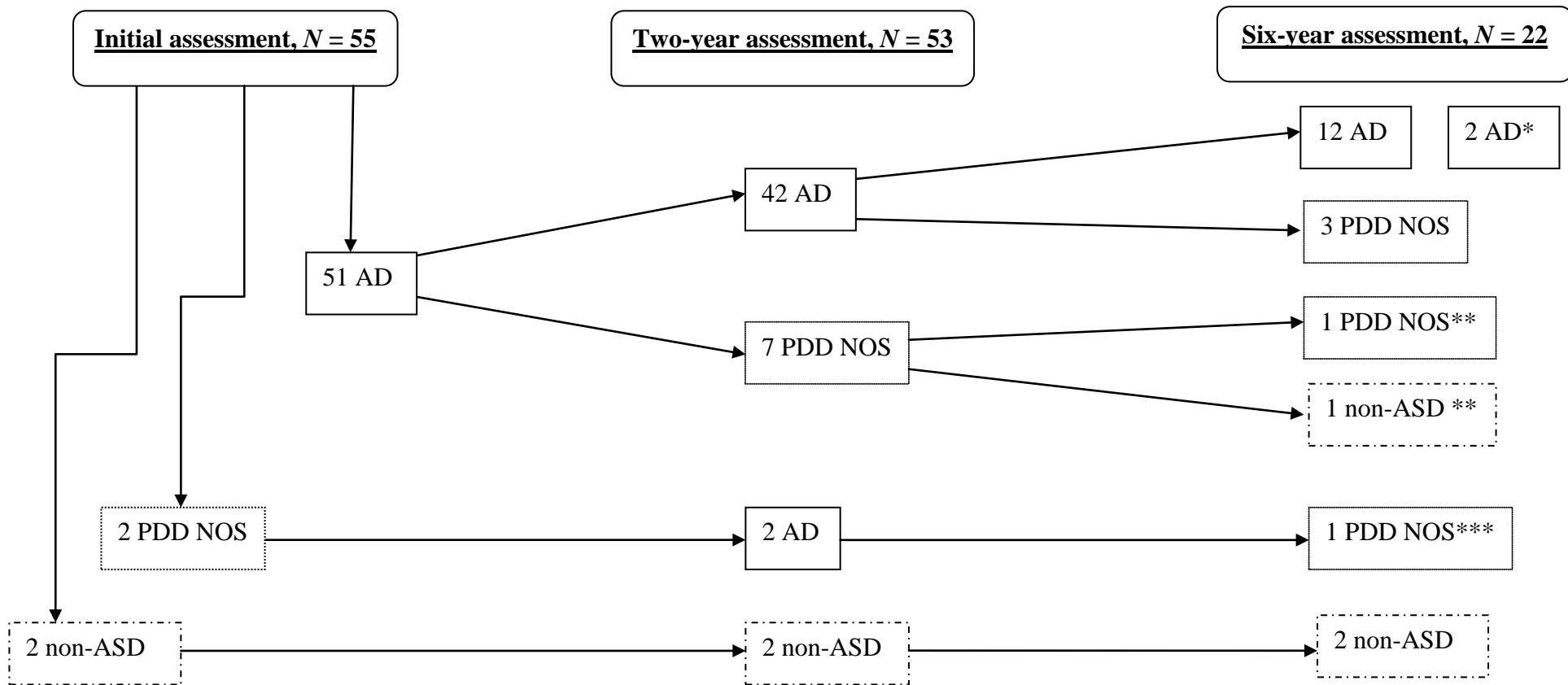SCQ score. The correlation matrix is presented in Table 9.

Figure 1. The change in diagnostic status of participants across the three time points.

Note:

* 2 AD at initial assessment did not attend the two-year follow up but still received AD classification at six-year follow up assessment.

** 5 PDD-NOS at two-year assessment did not attend the six-year follow up assessment.

*** 1 AD at two-year assessment did not attend the six-year follow up assessment.

Table 9

Pearson Correlation Analyses of ADEC and CARS with Other Outcome

Measures

Measure	Vineland	ADOS Rev	ADOS Severity	SCQ	WISC PRI	ABAS-II
	2 Years	6 Years	6 Years	6 Years	6 Years	6 Years
N	46	22	22	21	17	18
ADEC Initial	-.59**	.48*	.42	.44*	-.46	-.59*
CARS Initial	-.47**	.38	.39	.32	-.26	-.42

Note:
** $p < .01$ (2-tailed)
* $p < .05$ (2-tailed)

Past research examining predictors of long term outcomes in children

with ASD has pin-pointed IQ, especially non-verbal IQ, as one of the most

important predictors (Helt et al., 2008). However, measures of developmental or

intellectual functioning level of the participants were not available at the initial

assessment and, hence, adaptive functioning (specifically the Vineland ABC

score) was used as a control in the logistic regression models.

ADEC and CARS as Predictors of Two-year Follow Up

A binary hierarchical logistic regression was conducted to assess

whether the ADEC total score at initial assessment, after controlling for

participants' adaptive functioning level, significantly predicted whether

participants would be classified as AD based on DSM-IV-TR criteria at the two-

year follow up assessment. The full model containing the predictor (which was

the initial ADEC total score) was statistically significant, χ^2 (2, $N = 53$) = 22.65,

$p < .001$, $OR = 1.68$, 95% CI [1.15, 2.46], with the Hosmer and Lesmeshow test

indicating a good fit, χ^2 (8) = 4.18, p = .84. The model as a whole explained

between 35.3% (Cox and Snell R square) and 58.6% (Nagelkerke R square) of

the variance in DSM-IV-TR classification, and correctly classified 88.5% of

cases (as compared to 82.7% by chance). The odds of receiving an AD

classification on two-year follow up assessment increased between 1.15 to 2.46

times for every one-unit increase in the ADEC total score. For predicting the

DSM-5 ASD, the model was also statistically significant, χ^2 (2, N = 53) = 19.32,

p < .001, OR = 1.54, 95% CI [1.10, 2.15] and correctly classified 88.7% of cases

(as compared to 84.9% by chance).

    The full model containing the predictor (which was the initial CARS

total score) was also statistically significant, χ^2 (2, N = 53) = 12.53, p < .005, OR

= 1.36, 95% CI [1.06, 1.75], with the Hosmer and Lesmeshow test indicating a

good fit, χ^2 (8) = 5.99, p = .65. The model as a whole explained between 21.1%

(Cox and Snell R square) and 35.2% (Nagelkerke R square) of the variance in

DSM-IV-TR classification, and correctly classified 88.7% of cases (as compared

to 83% by chance). For predicting DSM-5 ASD, the model was also statistically

significant, χ^2 (2, N = 53) = 11.51, p < .005, OR = 1.32, 95% CI [1.03, 1.69] and

correctly classified 90.6% of cases (as compared to 84.9% by chance). The

considerable overlap of the 95% CIs for the odds ratios, suggests that both the

ADEC and CARS performed at a similar level.

    The full model containing both the predictors (which were the initial

ADEC and CARS score), again controlling for adaptive functioning, was also

explored. Results indicated that the model was significant, χ^2 (2, N = 53) =

23.76, p < .001, with the Hosmer and Lesmeshow test indicating a good fit, χ^2

(8) = 5.15, p = .74. The model as a whole explained between 36.1% (Cox and

Snell R square) and 60.4% (Nagelkerke R square) of the variance in DSM-IV-TR classification, and correctly classified 92.5% of cases (as compared to 83% by chance). Only the ADEC score became a significant individual predictor, $p < .05$, $OR = 1.61$, 95% CI [1.09, 2.39]. The assumption of no multicollinearity was considered violated if the tolerance value was less than 0.1 (Menard, 1995) and VIF value greater than 10 (Myers, 1990). In this case, the CARS' and ADEC's tolerance values were 0.52 and 0.68 respectively while their VIF values were 1.91 and 1.47 respectively. For predicting DSM-5 ASD, the model was also statistically significant, χ^2 (2, $N = 53$) $= 19.87$, $p < .001$ and correctly classified 94.3% of cases (as compared to 84.9% by chance). Similarly, only the ADEC was a significant individual predictor, $p < .05$, $OR = 1.48$, 95% CI [1.04, 2.10].

ADEC and CARS as Predictors of Six-year Follow Up

The full model was not statistically significant, χ^2 (2, $N = 22$) $= 4.49$, $p = .11$, $OR = 1.25$, 95% CI [.99, 1.58]. The model as a whole explained between 18.5% (Cox and Snell R square) and 25.3% (Nagelkerke R square) of the variance in DSM-IV-TR classification, and correctly classified 68.2% of cases (as compared to 63.6% by chance). When predicting DSM-5 ASD, the model was also not statistically significant, χ^2 (2, $N = 22$) $= 3.56$, $p = .17$, $OR = 1.22$, 95% CI [.95, 1.55].

Similar to the ADEC, the full model was not statistically significant, χ^2 (2, $N = 22$) $= 1.13$, $p = .57$, $OR = 1.13$, 95% CI [0.89, 1.44]. The model as a whole explained between 5% (Cox and Snell R square) and 6.8% (Nagelkerke R square) of the variance in DSM-IV-TR classification, and correctly classified 72.7% of cases (as compared to 63.6% by chance). Again, the model was also

not statistically significant for predicting DSM-5 ASD, χ^2 (2, $N = 22$) = 3.92, $p =$ .14, $OR = 1.31$, 95% CI [.94, 1.81].

Results indicated that the model was not significant, χ^2 (3, $N = 22$) = 4.52, $p = .25$. The model as a whole explained between 18.6% (Cox and Snell R square) and 25.4% (Nagelkerke R square) of the variance in DSM-IV-TR classification, and correctly classified 68.2% of cases (as compared to 63.6% by chance). Similarly, the model was also not statistically significant for predicting DSM-5 ASD, χ^2 (3, $N = 22$) = 4.82, $p = .19$ and correctly classified 81.8% of cases (as compared to 72.7% by chance).

Predictive Accuracy of ADEC and CARS

ROC analyses were used to investigate the ability of the ADEC and CARS score at initial assessment to predict DSM-IV-TR AD classification at two- and six-year follow up assessments. Using the area under the curve (AUC) as an index, an AUC can range from 0, a perfect negative classification, to .50, a completely chance outcome, to 1.0, a perfect prediction of AD using the ADEC. Interpretative guidelines propose that AUC values of .70 or above are considered moderate and .75 or above are considered good (Douglas, Guy, Reeves, & Weir, 2008).

The ADEC total score demonstrated good predictive accuracy at both the two-year follow up assessment, with an AUC of .92, 95% CI [.84, .99], and the six-year follow up assessment, with an AUC of .85, 95% CI [.69, 1.0] against DSM-IV-TR AD classification. The CARS had an AUC of .81, 95% CI [.64, .98] at two-year and an AUC of .72, 95% CI [.47, .98] at six-year follow up assessment against DSM-IV-TR AD classification. The CIs for the AUC values indicate acceptable predictive validity for DSM-IV-TR AD classification at the

two-year follow up assessments for both the ADEC and CARS, and for the

ADEC at the six-year assessment. However, the lower boundary of the CI for

the CARS at the six-year assessment indicates just below chance performance,

suggesting that further data would be valuable for assessing its predictive

accuracy at six years. ROC curves for both the ADEC and CARS at the two-year

and six-year follow up assessment are shown in Figure 2.

Discussion

This study extended prior research on the psychometric properties of two

Level 2 ASD screening tools by conducting the first examination of the

predictive validity of the ADEC and the CARS for long term outcomes such as

diagnostic classifications, symptom severity and functioning level in children

with ASD. The ADEC and CARS performed similarly in terms of predicting

clinical diagnostic outcome and overall adaptive functioning level at an interval

of two years following initial early childhood assessment. When both the CARS

and the ADEC scores were used together in the logistic regression models to

predict diagnostic outcome at two-year assessment following intervention, only

the ADEC score emerged as a significant individual predictor. However, given

the relatively small sample size for the six-year assessment, it is important that

these similarities and differences between the ADEC and the CARS are further

explored with new and larger samples. In the logistic regression models, neither
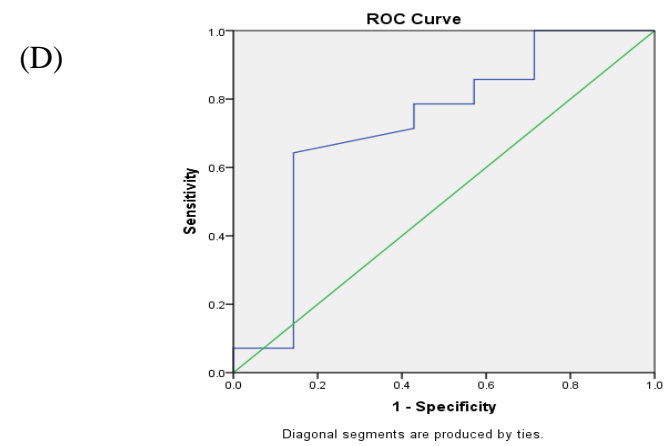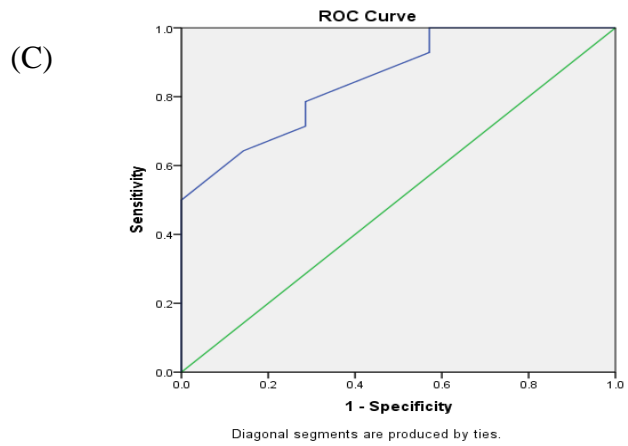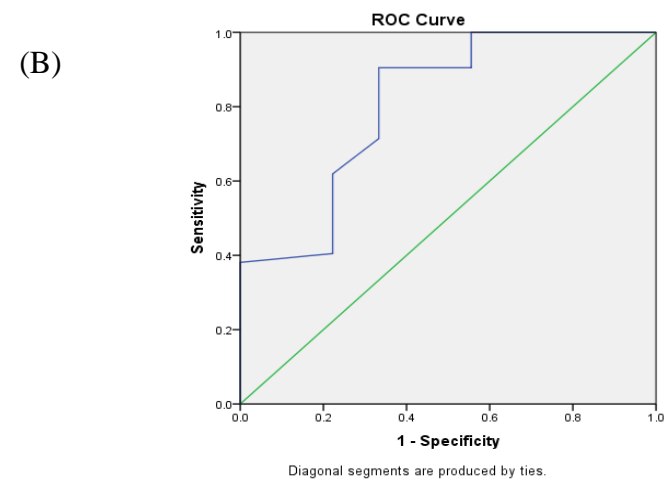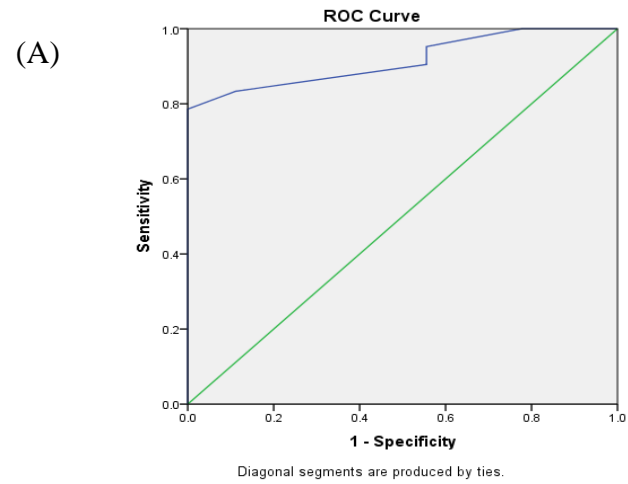
Figure 2. Receiver operating characteristic (ROC) curves.
Note: (A) ADEC for two-year follow up. (B) CARS for two-year follow up. (C) ADEC for six-year follow up. (D) CARS for six-year follow up.

the ADEC nor CARS score nor both ADEC and CARS scores accurately

predicted clinical diagnostic outcome at six-year follow up assessment. Yet,

using Kappa analysis, there was significant agreement between the initial ADEC

classification (but not the CARS) and the DSM-IV-TR diagnosis made at the

six-year assessment. In addition, only the ADEC score but not the CARS score

was found to be significantly correlated with ASD symptom severity (as

measured by ADOS revised algorithm total score and the SCQ score) measured

at the six-year assessment. However, neither the ADEC nor CARS score was

significantly correlated with ADOS severity score measured at the six-year

assessment.

In summary, this study is the first to report predictive validity for ADEC

and CARS early childhood assessments for long term outcomes such as AD

classification, ASD symptom severity, overall adaptive and/or cognitive

functioning taken at two- year and six-year follow up post initial assessments. In

addition, the predictive accuracy of the ADEC and CARS were also examined in

relation to the new DSM-5 criteria (APA, 2013). No other studies have been

conducted to examine the validity of using the Level 2 ASD screening tools to

predict long term outcomes in children with ASD.

Both the ADEC and the CARS demonstrated acceptable predictive

validity and suggested that ASD diagnosis can be stable, at least up to two years

post-initial assessment. As for the long term follow up (i.e., six-year), the

regression analysis did not predict AD classification using both the ADEC and

the CARS, though this finding could be affected by the smaller sample size at

the six-year follow up. However, when both the AD and the PDD-NOS were

combined into an ASD category, the majority of children who initially received

an ASD diagnosis continued to retain the ASD diagnosis at 2 years and 6 years

later. Nevertheless, differentiation within the autism spectrum appears to be

more challenging (Chawarska et al., 2007), and marked changes in clinical

presentation within the spectrum over time are to be expected as symptoms of

autism evolve and verbal and non-verbal cognitive skills improve (Chawarska et

al., 2009; Lord et al., 2006).

Limitations

There are some limitations of this study that should be considered. It

should be noted that the sample for the follow up assessments (especially for the

six-year data, $N = 22$) was small which increases the possibility of making a

Type II error. Hence, replication with larger samples together with age group

below 36 months at initial assessment should be carried out to evaluate the

generalizability of the present findings. In this study, the sample consisted

mostly of children with ASD and there is a need to include other populations of

children (especially more children with other developmental disorders) in future

predictive validity studies of ASD screening tools. With a larger sample, it will

also be important to consider covariates that may contribute to the predictive

validity of the ADEC (and the CARS).

Another possible limitation was that certain behaviours as reported in the

CARS were used to guide the decision making of the DSM-IV-TR ASD

classification in the sample. Results regarding consistency with diagnostic status

at the two-year follow up assessment might, therefore, need to be interpreted

with caution. However, the CARS individual item score or total score alone was

not relied on to make the diagnosis; rather, all available information from

developmental/psychological/medical assessment reports were used.

At the six-year follow up assessment, some children were evaluated in a university research centre and some were evaluated in their own homes. The use of two different settings may have resulted in some performance differences for the various assessment instruments (e.g., ADOS). However all available information (besides the ADOS) was used to guide the decision making of the clinical classifications of the participants. In addition, a post-hoc analysis showed that setting was not a significant predictor of whether participants met DSM-IV-TR and DSM-5 ASD classification.

Lastly, the focus of this study was to examine the ability of the CARS and ADEC to predict diagnostic classifications across time. It was not the aim in this study to speculate about child and family variables at initial assessment that might influence intervention outcome (e.g., a change in diagnostic classification from AD to PDD-NOS or PDD-NOS to non-ASD) or about why the participants did not want to return for the follow up studies. These issues should be examined in future research.

CHAPTER 4

Study 3: Development of a Brief Version of the Autism Detection in

Early Childhood (BADEC) for Ages from One to Six Years

This study reports the development of a brief version of the ADEC that would be suitable for clinicians to use in their practice. Psychometric properties of the ADEC have been reviewed and so far, the ADEC has been found to be a reliable and valid screening tool (Studies 1 & 2).

Currently, research by the Centers for Disease Control and Prevention indicates that only about 18% of children are diagnosed with an ASD by age 3 years (CDC, 2012), and more than half of children with developmental disabilities (including ASD) are not identified until they enter school, around age 4 years (Sices et al., 2003). Indeed, Shattuck et al. (2009) found that children could be identified with an ASD as late as 6 years old. Yet, there are few clinician-administered ASD screening instruments developed for children older than 36 months old. The ADOS (Lord et al., 2000) is one such clinician-administered tool that can be used for identifying older children with suspected ASD. It is typically used in the diagnostic process and has been shown to have excellent psychometric properties. However, its use requires specialized training, and it takes at least 45 minutes to administer. Screening for older children usually involves the use of parental rating scales such as the Social Communication Questionnaire (SCQ; Rutter et al., 2003) and the Social Responsiveness Scale/Social Responsiveness Scale 2nd Edition (SRS; Constantino & Gruber, 2006, SRS-2; Constantino & Gruber, 2012). Though parental rating scales have some significant advantages in terms of ease of

administration and reduced demands on professionals' time, there are concerns

with relying exclusively on parental reports, especially unstructured reports

(Barton et al., 2012). For instance, the rater may make subjective judgements

and it is difficult to monitor the accuracy of responses (Norris & Lecavalier,

2010). Previous work, primarily with older children and adolescents, found that

using data from multiple sources (i.e., clinicians, caregivers, and teachers)

enhances accuracy for the diagnosis of ASD (Kim & Lord, 2012a). For these

reasons it is important for clinicians to supplement the use of parent-report

instruments with clinician-administered screening tools. If the ADEC proved to

be a valid screening tool for ASD in older children it would meet the need for a

direct observational tool that can be administered by clinicians. Assessing the

ADEC's suitability as a screening tool for older children was one focus of this

study.

A second focus was ensuring the screening instrument was maximally

sensitive to any variations in symptoms observed in children of different ages. It

is clear that there are marked changes in clinical presentation within the

spectrum over time as symptoms of ASD evolve and verbal and non-verbal

cognitive skills develop (Lord et al., 2006; Chawarska et al., 2009). For

instance, a follow-up study found that some specific behaviours (such as

greeting, social reciprocity) which differentiated children with ASD from

children without ASD, were more prevalent when the children with ASD were 3

years old than at 2 years old (Lord, 1995). In another study using the ADI-R (Le

Couteur et al., 2003), features of children who were diagnosed with ASD at 42

months and those diagnosed with language disorder at 42 months were different

from when they were diagnosed at 20 months (Cox et al., 1999). For instance,

lack of pointing was one of the items that distinguished children with and without ASD at 20 months but not at 42 months and, similarly, lack of nodding differentiated children with and without ASD at 42 months but not at 20 months. Therefore, there may be certain behavioural markers of ASD that may be present in some younger children which we may not see in older children with ASD, and vice versa. In a similar vein, repetitive and stereotyped behaviours may be less obvious in younger children, although if these behaviours occur together with the social and communicative impairments, they are highly indicative of ASD (Charman et al., 2005). In other words, it seems likely that certain behaviours will be more salient and more predictive of ASD at certain ages. Precise information on the age at which different behavioural features are evident in children evaluated for ASD should highlight opportunities to improve the early detection of ASD.

In the first part of this study, I examined how the frequency and pattern of documented diagnostic features based on the ADEC varied by the age of the children with and without ASD. I examined the age groups 12-23 months, 24-35 months, 36-47 months, 48-59 months and 60-71 months and visually inspected whether there were any items for which there was a consistently low frequency of typical behaviours (such as responding to their name when called) displayed by children with ASD across the age groups. Failure to exhibit these typical behaviours may indicate pervasive autistic difficulties in children with ASD regardless of their age. I also examined whether there were any items of typical behaviour that increased in reported frequency in children with ASD with age. Such behaviour patterns may indicate that some older children with ASD have

learnt strategies to demonstrate these typical behaviours; alternatively, they may simply reflect maturation.

It is acknowledged that using 'visual inspection' may be subjective (Kazdin, 1994) and therefore I would augment the visual inspection with a ROC analysis of each ADEC item to determine which items are more salient and have better predictive value at certain ages.

To foreshadow my findings, the first part of the study identified behaviours with better diagnostic salience at some ages than others. Consequently, I developed a brief age-specific ADEC version for identifying children with possible ASD. The brief ADEC (BADEC) version should help clinicians and medical professionals confronted with children at different age levels to make a quick referral for a full diagnostic assessment.

A primary motivation for developing the brief versions of the ADEC arose from practitioners' reports that the currently available screening tools can take too long to administer (Barton et al., 2012). This has resulted in minimizing the uptake of ASD screening (AAP, 2003; Gura et al., 2011; Honigfeld & McKay 2006; Sices et al., 2003), with ASD screening rates ranging from 8% to 28% (Dosreis, Weiner, Johnson, & Newschaffer, 2006; Gillis, 2009). Although the ADEC consists of only 16 items and requires only 10-15 minutes to administer, paediatricians argue this may be prohibitive given their time constraint. In association with this study, I conducted a pilot survey of 30 medical practitioners who are practising in Australia to examine whether they would be keen to use a brief ASD screening tool in their practice. Results indicated that the majority of the respondents (N = 21; 70%) reported that they were likely to "use an ASD screening tool in their practice if the tool is easy to

administer and takes about 5 minutes". These survey responses reinforced the need to develop a time-efficient and age-specific version of the ADEC for clinicians to use in their settings. The brief version could consist of only 3 to 5 items which would ideally take less than 5 minutes to administer and to score. To emphasise, this brief ADEC version should not solely be used to make a one-off *diagnostic* decision (i.e., whether the child has ASD or not) but rather to help clinicians and medical professionals to make a quick *referral* for a full diagnostic assessment (i.e., whether the child should or should not be referred for an ASD assessment).

Study Aims

In sum, there were two main aims. First, I identified those behaviours tapped by ADEC items that were discriminating for each of five different age groups (12-23 months, 24-35 months, 36-47 months, 48-59 months and 60-71 months). Second, I developed a brief age-specific version of the ADEC (BADEC) for each of the different age groups. I examined the sensitivity, specificity, positive and negative predictive values associated with the different BADEC cutoff scores. I also calculated the internal consistency, diagnostic validity, concurrent validity and predictive validity for the BADEC version for each age group.

Method

Participants

The dataset included 457 participants, 197 of whom were involved in Study 1. Participants were aged 12–71 months of age, with the ethnic background of the sample being predominantly Caucasian (97.2%). Because these participants were not formally diagnosed at the time of screening, they

were not in any formal intervention programs. Data collection was spread over an 11-year period from 2003 to 2014, and the author was involved in the data collection[6] for this study from April 2011 to Oct 2014.

Of the 457 participants, 204 children had Autistic Disorder (AD), 55 children had pervasive developmental disorder - not otherwise specified (PDD-NOS), 20 children had Asperger's Disorder (AS), 91 children had other developmental disorders based on DSM-IV-TR diagnosis and 87 were considered typically developing. This study used archival and also prospective data collection to evaluate DSM-5 (APA, 2013) criteria with children with DSM-IV-TR clinical diagnoses. Informed consent was obtained from the participants' legal guardians and appropriate ethics approvals were also obtained prior to conducting this study.

A best estimate clinical (BEC) DSM-5 ASD diagnosis was made on each participant by the first author using all available information and assessment results, except for the ADEC data, to generate independent diagnoses. Participants were assigned to either one of three groups: (a) ASD; (b) non-ASD which include language and developmental delay, hearing loss and learning difficulty, and (c) typical development (TD). Sample characteristics (based on BEC DSM-5 ASD diagnosis) of all participants aged between 12 to 71 months are presented in Table 10.

Of the participants with DSM-5 ASD and non-ASD diagnoses in this study, inter-rater reliability was obtained for a subset of this sample of children (22%) using other independent professionals such as paediatricians and

---

[6] My role includes designing the study, recruiting participants, administering various tests such as the ADOS, ADI-R, Vineland and Mullen Scales to the participants, data analysis and write up.

psychologists who were recognized by the state's autism association to conduct ASD diagnostic assessments. The inter-rater reliability for DSM-5 diagnosis of ASD between the researcher's BEC diagnosis and the independent diagnosis was nearly perfect ($k = .97, p < .001$).

Diagnostic Evaluation Procedures and Measures

The procedures and materials used in this study have been previously described in Study 1. To summarise, parents and health care professionals who were concerned that their child or client presented with possible risk of developing an ASD participated in this screening study, and the children were assessed with a battery of tests such as the ADOS (Lord et al., 2000), ADI-R (Le Couteur et al., 2003), developmental and adaptive functioning assessments, where possible. The ADEC was also administered independently of the diagnostic assessment. To generate diagnoses independent of the ADEC, BEC DSM-IV-TR and DSM-5 diagnoses were made using all available information and assessment results. Because the ADOS – Toddler Module (ADOS-T; Lord et al., 2012) was released during the data collection phase of the present study, some of the more recent participants (specifically below 30 months old, $N = 20$) were administered the ADOS-T. The newly released ADI-R Toddler research algorithm (Kim & Lord, 2012b) was also used in my analyses.

Table 10

Description of Sample by Age Group and Diagnostic Classifications

Age Group		N	Gender (Male, Female)	Chronological age (months)	ADEC	ADOS Revised	ADOS Toddler	ADIR-T/ ADI-R	Non-Verbal IQ	VABC
										Measures
12-23 Months	ASD	22	19, 3	20.2 (2.3)	18.3 (5.8)	19.0 (6.1)	15.7 (5.2)	16.7 (4.9)	69.0 (18.6)	68.6 (9.7)
	Non-ASD	49	24, 25	17.7 (2.8)	6.7 (4.2)	8.8 (5.4)	4.8 (4.3)	10.0 (4.4)	86.0 (16.6)	79.2 (7.4)
	TD	43	24, 19	18.1 (2.2)	2.8 (2.2)	-	-	-	-	-
24-35 Months	ASD	72	58, 14	30.1 (3.3)	17.6 (6.3)	14.2 (5.1)	14.7 (5.8)	16.9 (5.6)	63.7 (16.8)	64.1 (8.4)
	Non-ASD	32	27, 5	29.3 (3.5)	5.9 (4.3)	4.3 (2.2)	6.0 (3.6)	5.8 (2.2)	75.5 (14.8)	73.6 (8.2)
	TD	29	14, 15	29.0 (3.8)	1.1 (1.7)	-	-	-	-	-
36-47 Months	ASD	93	70, 23	40.4 (3.5)	16.0 (6.7)	13.3 (5.6)	-	19.7 (6.8)	71.7 (27.5)	57.4 (10.2)
	Non-ASD	15	12, 3	39.7 (4.1)	6.9 (5.3)	6.2 (3.1)	-	11.1 (4.8)	70.9 (23.4)	69.4 (10.2)
	TD	15	9, 6	39.2 (3.5)	.00 (.00)	-	-	-	-	-
48-59 Months	ASD	45	37, 8	53.0 (3.3)	13.5 (7.9)	15.5 (5.0)	-	36.0 (11.2)	78.0 (10.1)	56.1 (13.6)
	Non-ASD	15	11, 4	53.2 (3.7)	3.9 (3.3)	6.1 (2.5)	-	29.1 (13.6)	69.3 (22.5)	68.0 (20.9)
60-71 Months	ASD	19	11, 8	64.4 (4.1)	10.1 (8.7)	11.5 (2.8)	-	33.1 (8.7)	90.5 (14.7)	59.2 (17.6)
	Non-ASD	8	5, 3	64.1 (3.9)	.8 (1.5)	5.8 (5.4)	-	14.8 (11.1)	90.6 (8.2)	67.0 (11.3)

Note: ADOS Revised Autism Diagnostic Observation Schedule – Revised total algorithm score, *ADIR-T* Autism Diagnostic Interview Revised Toddler total algorithm score, *IQ* Mullen Scales of Early Learning or Bayley-III intellectual quotient, *VABC* Vineland Adaptive Behavior Composite standard score.

ADOS-Toddler. The ADOS has been deemed to be more accurate and valid than the ADI-R when used in diagnosing children under the age of three (Gray, Tonge, & Sweeney, 2008; Chawarska et al., 2007), but research has indicated that it remains of limited value for children with non-verbal mental ages below 16 months (Gotham et al., 2007). For this young population, the ADOS Module 1 algorithm tends to over- classify about 81% (19% specificity) of children with intellectual disabilities and/or language impairments as having autism or ASD when clinical judgement deems that they do not.

As a result, the ADOS – Toddler Module (ADOS-T; Lord et al., 2012) was developed to further improve sensitivity and specificity of the ADOS as a diagnostic instrument for children under 30 months of age who have minimal speech (ranging from no spoken words to simple two-word phrases), have a non-verbal age equivalent of at least 12 months and are walking independently. The Toddler module follows the same structure as other modules of the ADOS, with the examiner presenting semi-structured and motivating activities for the child and observing the child's responses as well as attempts to maintain the interaction. Symptoms relevant to a diagnosis of ASD are scored from 0 to 3 on the ADOS-T, with higher numbers indicating more abnormality. Two diagnostic algorithms have been derived for the ADOS-T: (1) for children between 12 and 20 months and non-verbal children who are 21–30 months, and (2) for verbal children between 21 and 30 months. Diagnostic algorithms yield domain scores for Social Affect (SA) and Restricted and Repetitive Behaviours (RRB) and a total algorithm score (Luyster et al., 2009). Based on cutoffs applied to total scores, the ADOS-T yields just two classifications: ASD and Nonspectrum. The ADOS-T also yields three ranges of concern: 'little-to-no concern,' 'mild-to-

moderate concern,' and 'moderate-to-severe concern'. The algorithm scores have acceptable internal consistency and excellent inter-rater and test–retest reliability. The algorithm, using both the formal cutoff and the ranges of concern, has excellent diagnostic validity for ASD versus non-autism spectrum disorders.

ADIR-Toddler. Because of the criticism of the original ADI-R with respect to diagnosing toddlers, a toddler version of the ADI-R has been developed and in use that contains additional questions relating to early childhood behaviours; however, the scoring criteria remain the same as for the standard ADI-R (Bishop, Luyster, Richler, & Lord, 2008; Lord et al., 2004). Hence, the new ADI-R algorithms (Kim & Lord, 2012b) were developed to extend the valid use of the ADI-R to toddlers and young preschoolers ranging from 12 to 47 months and down to non-verbal mental age of 10 months. Using the new algorithms for toddlers and preschool children led to improved sensitivity and specificity as compared to the originally developed algorithm (Kim & Lord; Kim, Thurm, Shumway & Lord, 2012). Similar to the ADOS-T, the ADIR-Toddler (ADIR-T) provides diagnostic algorithms scores for Social Affect (SA) and Restricted and Repetitive Behaviours (RRB) domains (which are based on the new DSM-5 criteria) and a total algorithm score and with just two classifications: ASD and Nonspectrum. In addition, the ADIR-T uses the ranges of concern concept (i.e., 'little-to-no concern,' 'mild-to-moderate concern,' and 'moderate-to-severe concern'). For children aged 48 months and above, I used the original ADI-R algorithm.

DSM-5 Criteria. In order to help to guide the decision-making of the BEC DSM-5 diagnosis, and in addition to using all available information and

assessment results, I also relied on the supplementary tables provided by Huerta et al. (2012) where items from the ADOS or/and the ADI-R are mapped onto the DSM-5 criteria. DSM-5 guidelines were then followed to determine whether each participant met or did not meet the DSM-5 criteria for ASD.

Results

Examination of Typical Behaviours on Each ADEC Item by Age Group

First, I examined the frequency and pattern of children with and without ASD who displayed typical behaviours (as defined by a score of 0, i.e., the child performs the expected or example behaviour as operationalised in the manual) on the ADEC items across the different age groups (refer to Figure 3). Based on visual inspection of the percentage of participants showing typical behaviour on each ADEC item, children with ASD showed a lower frequency of typical behaviours as compared to children without ASD across all the age groups. Next, I examined the individual ADEC items in greater detail. On the one hand, I noted that there were some items (e.g., items 1, 4, 5, 8 and 15) where there was a consistently low frequency of typical behaviours (about 20-40%) displayed by children with ASD across most of the age groups. These behaviours may indicate pervasive failure to learn or develop these typical skills in the children with ASD. On the other hand, there were other items (e.g., items 9, 13 and 14) where there was a consistently higher frequency of typical behaviours (averaging about 60%) displayed by children with ASD across the age groups. Lastly, there were some items (e.g., items 2, 3, 6, 7, 10, 11 and 12) where there was a trend of an increasing frequency of typical behaviours observed in children with ASD across the age groups, which may indicate learned behaviour or simply effects of maturation.

From Figure 3, it appears that there may be certain behaviours that are more salient and have better predictive value at certain ages. Of particular note, there were some key behaviours (i.e., response to name, gaze switch, eye contact, social smile and use of gestures) which children with ASD infrequently demonstrated, regardless of their age.With this information, my second aim of the study was to examine whether I could identify a brief age-specific version of the ADEC (BADEC) for the different age groups using some of these key behaviours identified on the ADEC.

Data Analysis

For each age group of participants, I examined the initial area under the curve (AUC) value for each of the ADEC items. Recall that, the area under the curve (AUC) is a measure of the overall predictive validity, where an AUC = .5 indicates random prediction of the independent variable and an AUC of > .75 indicates good validity (Douglas et al., 2008). In the event where several ADEC items were presented with the same AUC value, I would then choose the item with the narrower confidence interval. I then compared which version, either using the 5 key items identified in Figure 3 (i.e., 1, 4, 5, 8 and 15) or the 5 highest AUC items, to derive the brief ADEC version (BADEC) for each age group, and compared it with the AUC for the full versions.
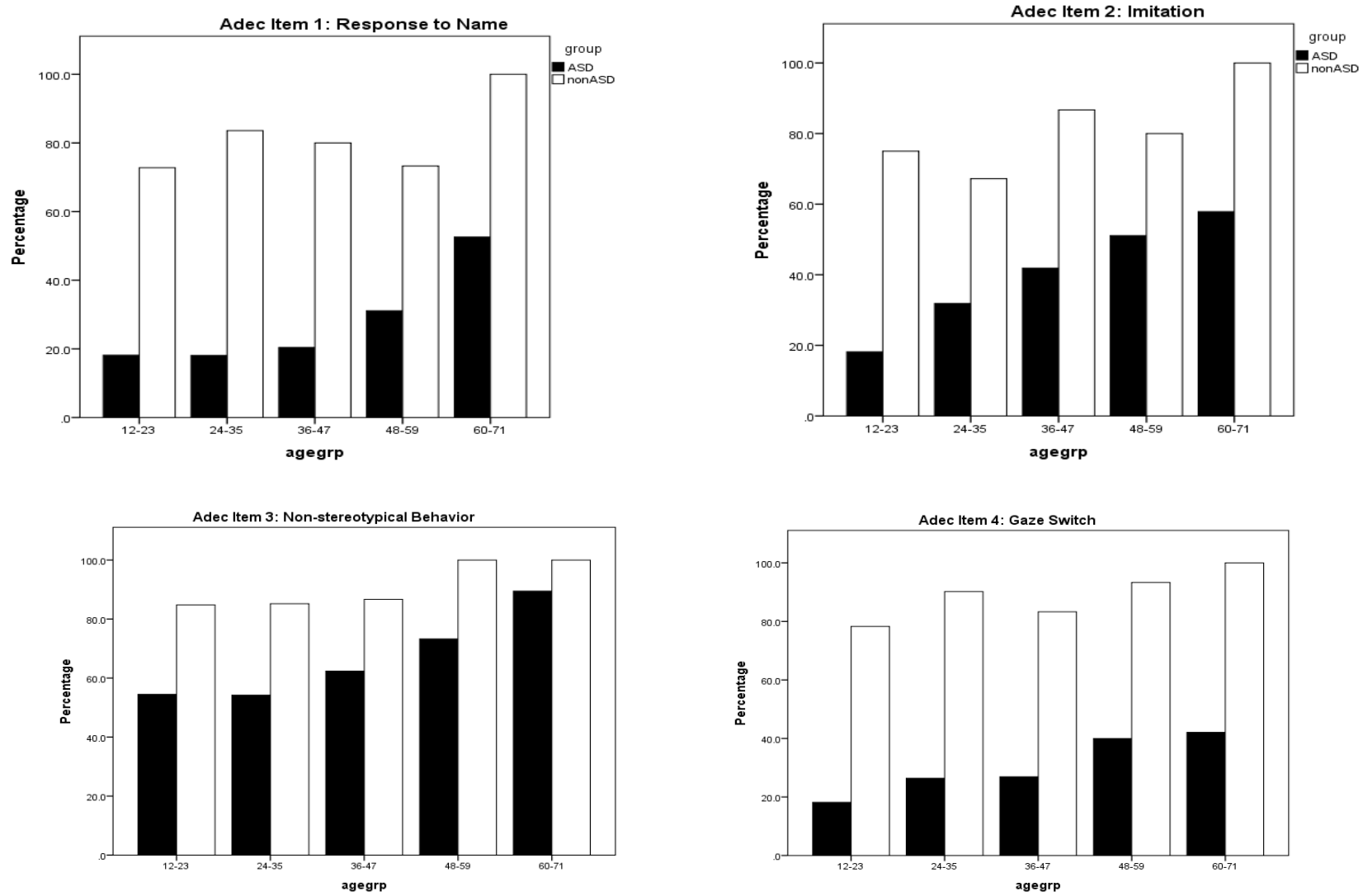
Figure 3. Patterns of typical behaviours by children with and without ASD on each ADEC item across the age groups.
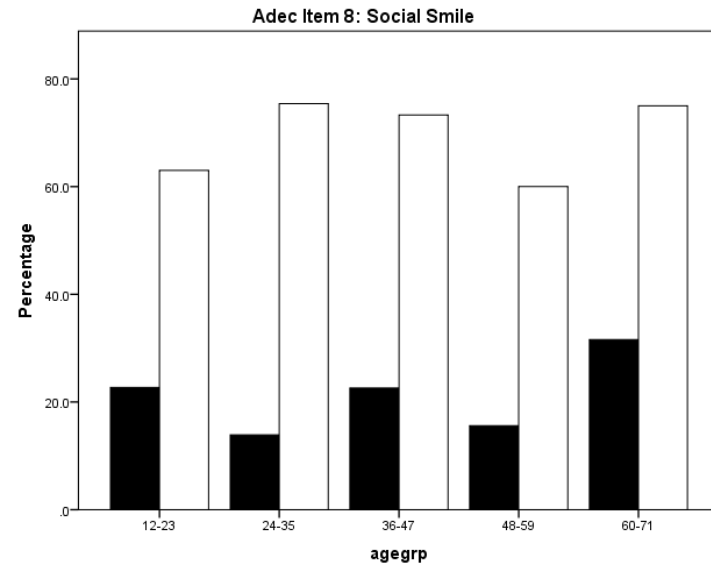
Figure 3. Patterns of typical behaviours by children with and without ASD on each ADEC item across the age groups (continued).

Figure 3. Patterns of typical behaviours by children with and without ASD on each ADEC item across the age groups (continued).
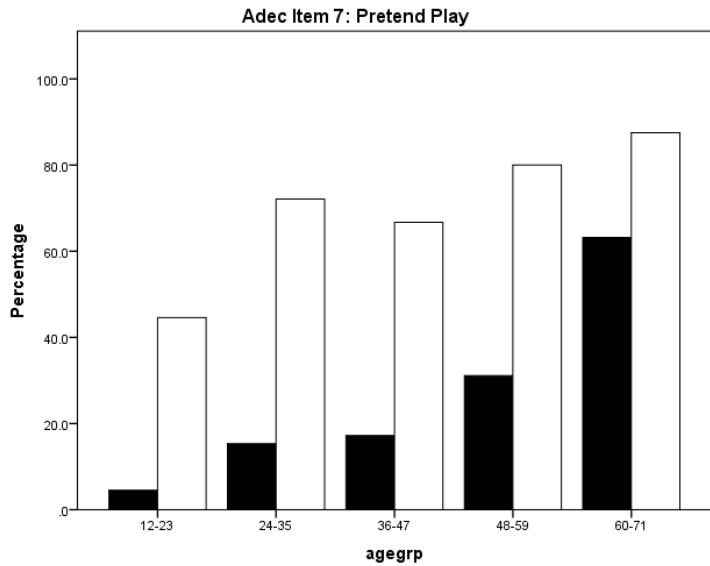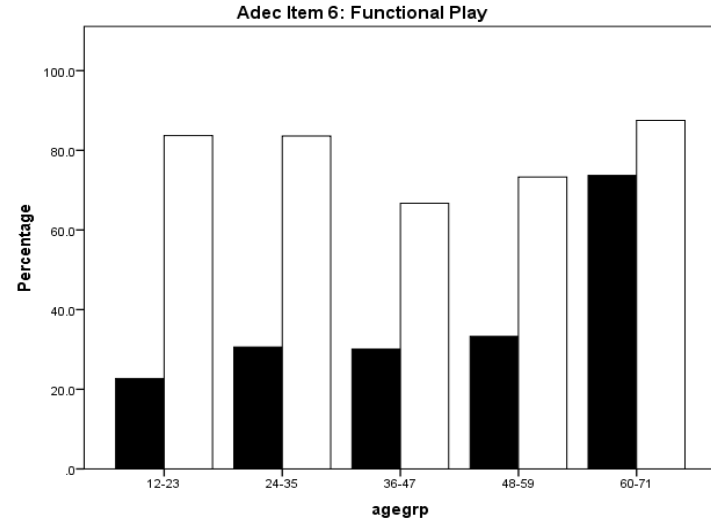
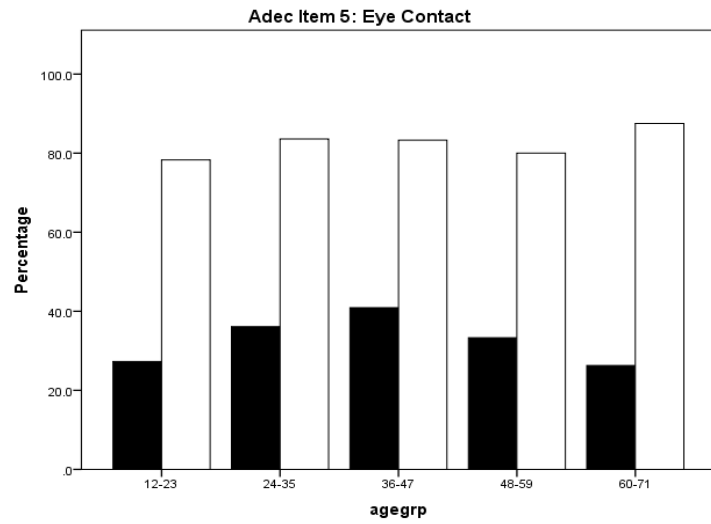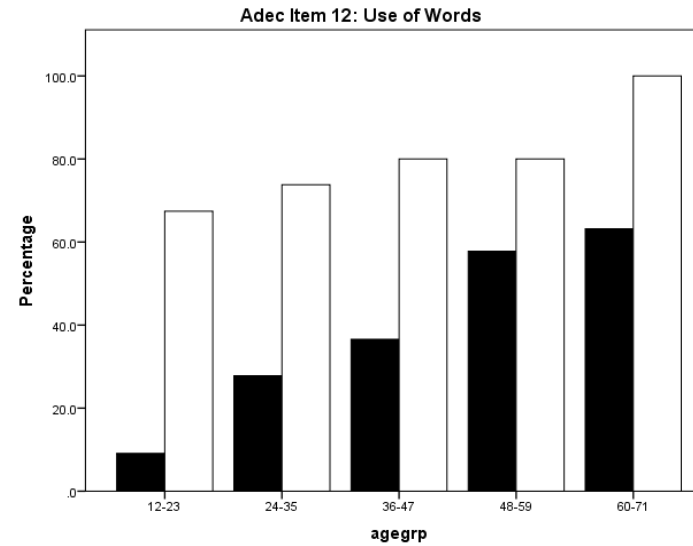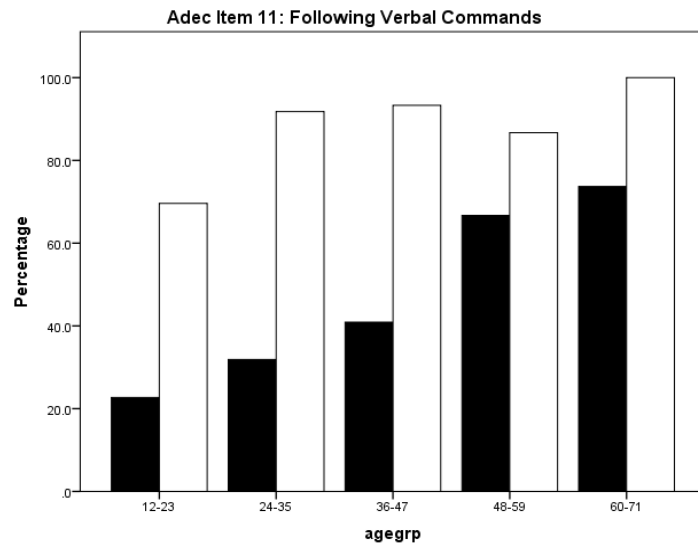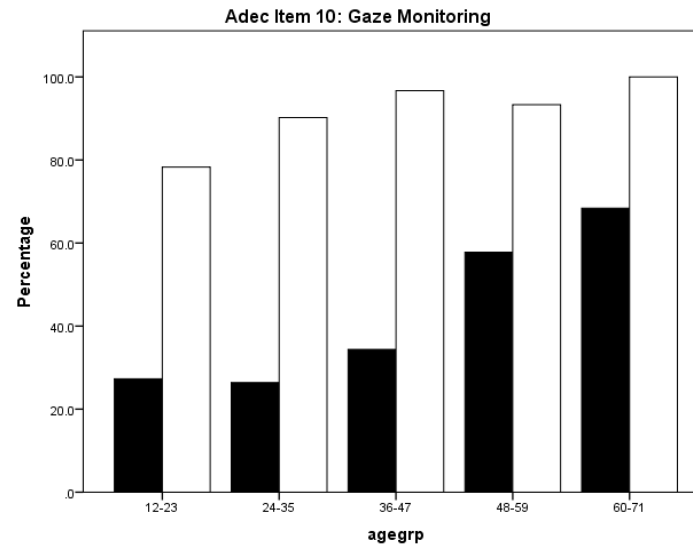*Figure 3.* Patterns of typical behaviours by children with and without ASD on each ADEC item across the age groups (continued).
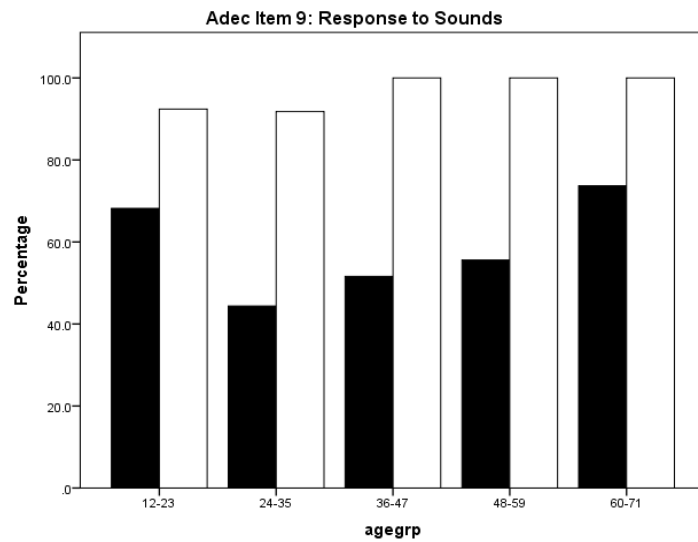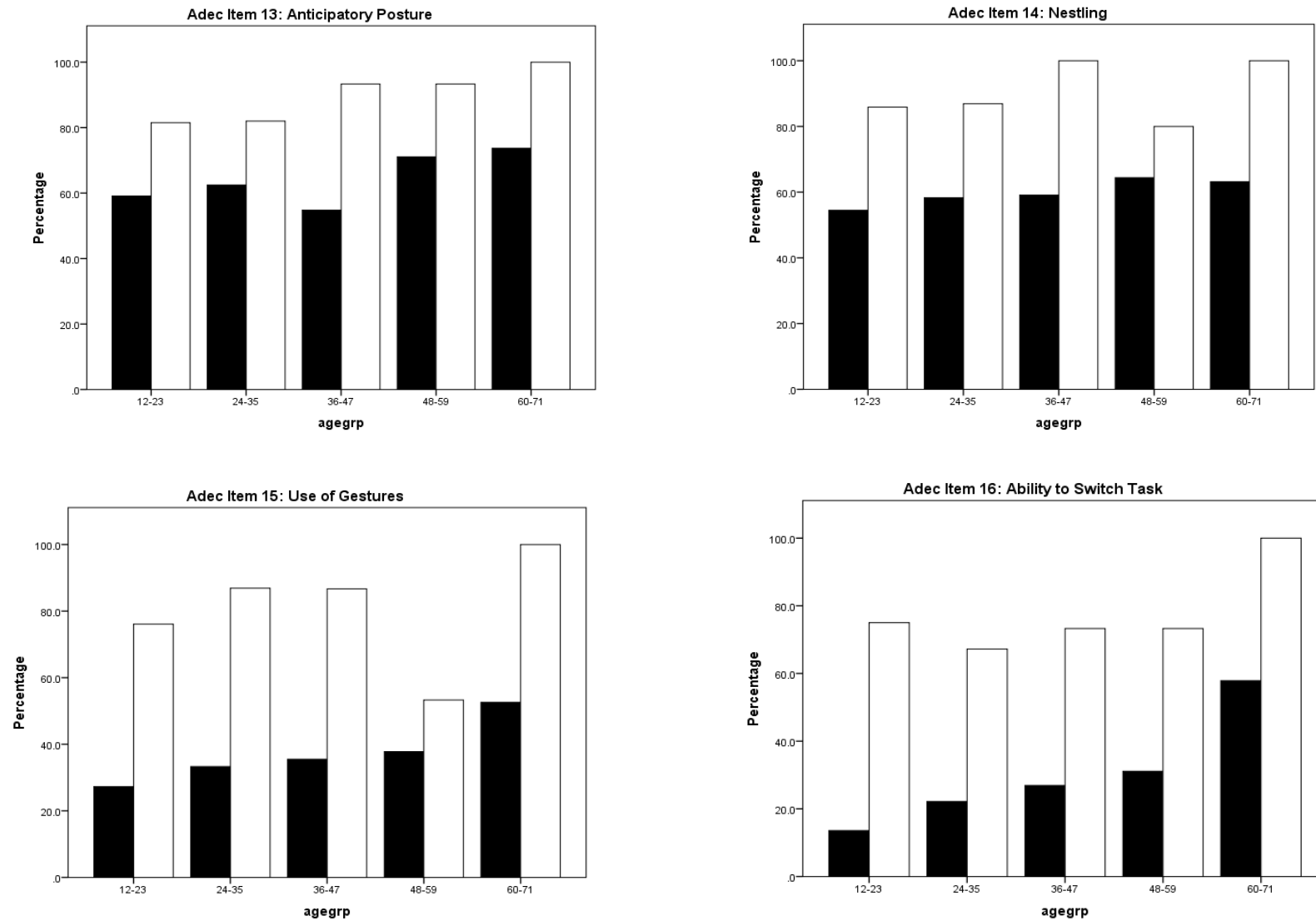
It has been argued that routine developmental surveillance should be performed at all primary care practice visits of well children (and also during sick visits to sick children and immunisations) from infancy through school-age in (Filipek et al., 2000), where medical professionals evaluate *all* children (whether typical developing or with suspected delay) for any possible developmental difficulties. It has been proposed that medical professionals may benefit from using standardised screening tools during the developmental surveillance to aid in their decision-making (Miller et al., 2011). To simulate the process of developmental surveillance and screening in clinical setting, for my younger groups (47 months and below), analyses were conducted using children in the ASD group versus the non-ASD combined with TD groups. I did not have any TD children in the older age groups (48 months and above), perhaps because most of the developmental concerns (such as speech delay) had already been identified or resolved in the older children when they enter school at around 4 years of age (Sices et al., 2003). So for the older age groups (48 months and above), analyses were done using ASD and non-ASD samples.

I computed the sensitivity, specificity, positive and negative predictive values associated with the different BADEC versions' cutoff scores, and calculated Cronbach's alpha for each brief version. Diagnostic validity analyses using ANOVA were conducted separately for (1) < 47 months children in the (a) ASD versus non-ASD groups and (b) ASD versus non-ASD combined with TD groups and (2) > 48 months children in the ASD versus non-ASD groups. Correlations were examined between total scores on the brief and full version of the ADEC, and also with the ADOS revised algorithm score, ADOS-T total score (where available) and ADIR-T/ADI-R total score. Finally, I compared the

predictive validity of the BADEC scores on DSM-5 classification of ASD against the ADOS revised algorithm score and ADIR-T/ADI-R total score using a binary logistic regression analysis.

ROC Analysis of ADEC Items for All Age Groups

First, I examined the mean (and standard deviation), together with the AUC of each ADEC item (95% confidence interval, CI) (Table 11). When I examined the 5 ADEC items with the highest AUC in each age group, there were 2 ADEC items (i.e., items 1: Response to name and 4: Gaze switch) that were present across all the age groups; item 8: Social smile, appeared in all age groups except for the 12-23 months group; item 5: Eye contact, appeared in the 48-59 and 60-71 months groups and item 15: Use of gestures, appeared in the age groups of 24-35 and 60-71months old. I then investigated whether I could use these 5 items to form *one* brief version for all age levels instead of developing separate forms based on different age groups. The obvious rationale for using one form instead of separate forms would be to facilitate the clinicians' ease of use. Nevertheless, I also compared the AUC value of using these 5 key items (i.e., 1, 4, 5, 8 and 15) to see whether its AUC would be sufficiently larger than using the AUC value of the 5 best items (as defined by the highest AUC and thus might or might not be the same as the 5 key items) in each age group. If the 5 best items version produced a higher AUC (as defined by no overlap in the 95% confidence interval) than the 5 key items version, the 5 best items version would then be chosen as the brief ADEC (BADEC) version for each age group. I then compared this BADEC version to the full ADEC version for each age group.

I found that while the 5 best items produced higher AUC values than the 5 key items, the AUC confidence intervals overlapped (see Table 12). Hence, there appeared to be no meaningful difference between using the 5 key items versus using the 5 best items in each age group. Consequently, we used the 5 key ADEC items to form *one* version for all the age groups. Cronbach's alpha values for the chosen brief versions are also presented in Table 12.

Next, I examined the optimal cutoff score for the different age versions (Table 13). Generally, the cutoff score became lower as the age increased, with sensitivity ranging from .82 to .90 and specificity ranging from .53 to .89. It was noted that using a cutoff score of 2, specificity was low (.53) for the 48-59 months group because I wanted to maximise the sensitivity (.89). Clinicians could also use the cutoff score of 3 if the low specificity was an issue, though both sensitivity and specificity were also low using this cutoff of 3.

Table 11

Mean, Standard Deviation and AUC Values for Each ADEC Item for Each Age Group

ADEC Item	12-23 Months			24-35 Months		
	Mean (SD)		AUC (95% CI)	Mean (SD)		AUC (95% CI)
	<u>ASD</u> (n=22)	non-<u>ASD</u>+TD (n=92)		<u>ASD</u> (n=72)	non-<u>ASD</u>+TD (n=61)	
1. Response to name[a][b]	1.41 (.80)	.33 (.58)	**.83 (.72, .94)**	1.36 (.78)	.16 (.37)	**.87 (.81, .93)**
2. Imitation	.95 (.58)	.32 (.59)	.77 (.67, .88)	1.13 (.87)	.43 (.67)	.72 (.63, .80)
3. Ritualistic play	.64 (.79)	.15 (.36)	.67 (.52, .81)	.61 (.74)	.16 (.42)	.66 (.57, .76)
4. Joint attention and social referencing[a][b]	1.32 (.78)	.23 (.45)	**.85 (.74, .96)**	1.26 (.86)	.11 (.37)	**.84 (.77, .91)**
5. Eye contact	1.09 (.81)	.23 (.45)	.79 (.67, .91)	.90 (.79)	.20 (.48)	.75 (.67, .83)
6. Functional play[a]	1.18 (.73)	.20 (.47)	**.82 (.71, .93)**	1.15 (.87)	.20 (.48)	.79 (.71, .87)
7. Pretend play	1.70 (.59)	.86 (.86)	.77 (.67, .87)	1.57 (.75)	.43 (.74)	.82 (.75, .90)
8. Reciprocity of smile[b]	1.27 (.83)	.39 (.53)	.79 (.66, .91)	1.49 (.73)	.33 (.63)	**.85 (.78, .92)**
9. Reaction to common sounds	.43 (.66)	.09 (.32)	.62 (.48, .77)	.81 (.82)	.10 (.35)	.74 (.66, .83)
10. Gaze monitoring[b]	1.36 (.90)	.30 (.62)	.79 (.67, .91)	1.28 (.86)	.13 (.43)	**.83 (.76, .91)**
11. Following verbal commands	1.16 (.75)	.34 (.54)	.78 (.66, .90)	1.14 (.88)	.10 (.35)	.81 (.74, .89)
12. Delayed language[a]	1.73 (.63)	.43 (.68)	**.88 (.79, .96)**	1.31 (.88)	.48 (.83)	.73 (.64, .82)
13. Anticipation of social advances	.73 (.94)	.24 (.54)	.63 (.49, .77)	.57 (.80)	.21 (.49)	.61 (.51, .70)
14. Nestling	.73 (.88)	.14 (.35)	.68 (.53, .82)	.58 (.77)	.15 (.40)	.65 (.56, .74)
15. Use of gestures[b]	1.41 (.91)	.35 (.67)	.79 (.66, .91)	1.24 (.93)	.13 (.34)	**.81 (.73, .88)**
16. Task switching[a]	1.14 (.64)	.26 (.47)	**.84 (.74, .93)**	1.18 (.78)	.34 (.51)	.79 (.71, .86)

Note. The 5 ADEC items with the highest AUC are in bold.

[a] The 5 ADEC items for 12-23 Months Group.

[b] The 5 ADEC items for 24-35 Months Group.

Table 11 (Continued)

Mean, Standard Deviation and AUC Values for Each ADEC Item for Each Age Group

ADEC Item	36-47 Months			48-59 Months			60-71 Months		
	Mean (SD)		AUC	Mean (SD)		AUC	Mean (SD)		AUC
	ASD	non-ASD+TD	(95% CI)	ASD	non-ASD	(95% CI)	ASD	non-ASD	(95% CI)
	(n=93)	(n=30)		(n=45)	(n=15)		(n=19)	(n=8)	
1. Response to name[cde]	1.15 (.74)	.27 (.58)	**.81 (.72, .90)**	1.07 (.84)	.27 (.46)	**.76 (.64, .88)**	.68 (.82)	.00 (.00)	**.74 (.55, .92)**
2. Imitation	.95 (.89)	.20 (.55)	.73 (.63, .82)	.78 (.88)	.33 (.72)	.64 (.48, .80)	.79 (.98)	.00 (.00)	.71 (.52, .90)
3. Ritualistic play	.48 (.69)	.13 (.35)	.63 (.52, .73)	.38 (.68)	.00 (.00)	.63 (.49, .78)	.11 (.32)	.00 (.00)	.55 (.32, .79)
4. Joint attention and social referencing[cde]	1.18 (.83)	.23 (.57)	**.80 (.71, .88)**	1.07 (.94)	.07 (.26)	**.78 (.67, .90)**	.95 (.91)	.00 (.00)	**.79 (.62, .96)**
5. Eye contact[de]	.85 (.81)	.17 (.38)	.73 (.64, .82)	.89 (.75)	.20 (.41)	**.76 (.63, .88)**	1.05 (.78)	.13 (.35)	**.83 (.67, .99)**
6. Functional play	1.05 (.81)	.47 (.73)	.70 (.59, .80)	1.04 (.85)	.33 (.62)	.73 (.59, .87)	.47 (.84)	.13 (.35)	.58 (.36, .81)
7. Pretend play[d]	1.54 (.77)	.60 (.89)	.76 (.65, .86)	1.36 (.93)	.33 (.72)	**.77 (.63, .90)**	.68 (.95)	.25 (.71)	.62 (.39, .84)
8. Reciprocity of smile[cde]	1.31 (.82)	.33 (.61)	**.80 (.71, .89)**	1.33 (.74)	.47 (.64)	**.79 (.67, .92)**	1.16 (.90)	.25 (.46)	**.78 (.60, .95)**
9. Reaction to common sounds	.67 (.77)	.00 (.00)	.74 (.66, .83)	.53 (.66)	.00 (.00)	.72 (.60, .85)	.26 (.45)	.00 (.00)	.63 (.42, .85)
10. Gaze monitoring[c]	1.14 (.90)	.03 (.18)	**.82 (.75, .89)**	.73 (.92)	.07 (.26)	.69 (.55, .82)	.53 (.84)	.00 (.00)	.66 (.48, .89)
11. Following verbal commands[c]	1.00 (.91)	.10 (.40)	**.77 (.68, .85)**	.53 (.82)	.13 (.35)	.61 (.46, .76)	.37 (.68)	.00 (.00)	.63 (.42, .85)
12. Delayed language	1.09 (.91)	.33 (.71)	.72 (.62, .82)	.78 (.95)	.33 (.72)	.62 (.46, .77)	.63 (.90)	.00 (.00)	.68 (.48, .89)
13. Anticipation of social advances	.74 (.88)	.10 (.40)	.70 (.60, .79)	.42 (.72)	.07 (.26)	.62 (.47, .77)	.42 (.77)	.00 (.00)	.63 (.42, .85)
14. Nestling	.59 (.78)	.00 (.00)	.70 (.61, .80)	.47 (.69)	.27 (.59)	.58 (.42, .74)	.53 (.77)	.00 (.00)	.68 (.48, .89)
15. Use of gestures[e]	1.14 (.92)	.20 (.55)	.77 (.68, .86)	1.13 (.94)	.73 (.88)	.61 (.45, .77)	.84 (.96)	.00 (.00)	**.74 (.55, .92)**
16. Task switching	1.09 (.79)	.30 (.54)	.77 (.68, .86)	.96 (.77)	.33 (.62)	.72 (.58, .87)	.63 (.83)	.00 (.00)	.71 (.52, .90)

Note. The 5 ADEC items with the highest AUC are in bold.
[c] The 5 ADEC items for 36-47 Months Group. [d] The 5 ADEC items for 48-59 Months Group. [e] The 5 ADEC items for 60-71 Months Group.

Table 12

AUC Values (with 95% CI) for the Different ADEC Versions for Each Age

Group and Cronbach's Alpha for Chosen Version (in bold)

Age Group	5 Key ADEC Items	5 Highest ADEC Items	Full ADEC Items	Cronbach's Alpha
12-23 Months	**.93 [.88, .99]**	.97 [.95, .99]	.97 [.95, .99]	.82
24-35 Months	**.95 [.92, .99]**	.96 [.93, .99]	.97 [.94, .99]	.86
36-47 Months	**.91 [.85, .97]**	.93 [.88, .99]	.93 [.87, .99]	.77
48-59 Months	**.82 [.72, .93]**	.87 [.79, .96]	.84 [.75, .94]	.86
60-71 Months	**.87 [.74, .99]**	.87 [.74, .99]	.88 [.76, 1.00]	.89

ANOVA Analysis of BADEC Versions for All Age Groups

As indicated in Table 14, diagnostic validity of the different BADEC

versions was established for all the age group versions, indicating that the

BADEC versions were able to differentiate between the ASD and non-ASD

groups. An analysis of covariance (ANCOVA) with the non-verbal IQ specified

as a covariate was performed in the 12-23 and 24-35 months group due to the

significant difference in non-verbal IQ between the participants with and

without ASD while an ANOVA was performed for the other age groups.

Table 13

Sensitivity and Specificity Associated with the Different BADEC Cutoff Score

BADEC Score	12-23 Months				24-35 Months				36-47 Months				48-59 Months				60-71 Months			
	Sen	Spe	PPV	NPV	Sen	Spe	PPV	NPV	Sen	Spe	PPV	NPV	Sen	Spe	PPV	NPV	Sen	Spe	PPV	NPV
1	1.0	.33	.26	1.0	.99	.54	.72	.97	.99	.63	.89	.95	.93	.27	.79	.57	**.84**	**.75**	**.89**	**.67**
2	1.0	.58	.36	1.0	.97	.82	.86	.96	.94	.73	.92	.79	**.89**	**.53**	**.85**	**.62**	.74	.88	.93	.58
3	.86	.79	.50	.96	**.90**	**.87**	**.89**	**.88**	**.86**	**.83**	**.94**	**.66**	.69	.67	.86	.42	.58	1.0	1.0	.50
4	**.82**	**.89**	**.64**	**.95**	.81	.93	.94	.80	.75	.83	.93	.52	.69	.80	.91	.46	.58	1.0	1.0	.50
5	.73	.91	.67	.93	.74	.97	.96	.76	.65	.83	.92	.43	.58	1.0	1.0	.44	.58	1.0	1.0	.50
6	.68	.96	.79	.93	.64	.97	.96	.69	.58	.97	.98	.43	.49	1.0	1.0	.40	.53	1.0	1.0	.47
7	.50	.99	.92	.89	.49	.98	.97	.62	.45	.97	.98	.36	.44	1.0	1.0	.38	.42	1.0	1.0	.42
9	.32	1.0	1.0	.86	.28	1.0	1.0	.54	.12	1.0	1.0	.27	.27	1.0	1.0	.31	.16	1.0	1.0	.33
10	.18	1.0	1.0	.84	.08	1.0	1.0	.48	.05	1.0	1.0	.25	.16	1.0	1.0	.28	.05	1.0	1.0	.31

Note: Sen = Sensitivity, *Spe* = Specificity, *PPV*=Positive Predictive Value, *NPV*=Negative Predictive Value

Table 14

ANOVA/ANCOVA Analyses for the BADEC Versions for Each Age Group

Age Group	Groups	Mean	df	F	$\eta2$
		BADEC (SD)			
12-23 Months	ASD	6.0 (3.1)	1, 48	13.27**	.22
	Non-ASD	2.3 (1.9)			
24-35 Months	ASD	5.4 (2.6)	1, 34	26.11**	.43
	Non-ASD	1.3 (1.1)			
36-47 Months	ASD	5.6 (2.5)	1, 106	21.15**	.17
	Non-ASD	2.4 (2.4)			
	Non-ASD+	1.2 (2.0)	1, 121	74.92**	.38
	TD				
48-59 Months	ASD	5.5 (3.3)	1, 58	17.56**	.23
	Non-ASD	1.7 (1.5)			
60-71 Months	ASD	4.7 (3.5)	1, 25	11.88*	.32
	Non-ASD	.4 (.7)			

Note. *p < .005 **p < .001

Correlation Analysis of BADEC Versions for All Age Groups

Results (in Table 15) indicated that all the BADEC versions correlated

significantly with the ADEC full versions (rs= .88 to 95) and the ADOS revised

algorithm scores (rs= .68 to 88, with the exception of the 48-59 months group).

Both the 12-23 months and 24-35 months BADEC version also correlated with

the ADOS Toddler total score (rs= .69 to 93). The BADEC score was only

significantly correlated with the ADIR-T/ADI-R score for age groups of 24-35 and 60-71 months old.

Table 15

Correlation of the Different BADEC Version Scores with Various Measures

BADEC Version for Age Groups	ADEC Full Version	ADOS Rev	ADOS-T	ADIR-T/ ADI-R
12-23 Months (*N*)	.92**(*114*)	.83**(*15*)	.69*(*10*)	.38(*17*)
24-35 Months (*N*)	.94**(*133*)	.88**(*25*)	.93**(*9*)	.38*(*36*)
36-47 Months (*N*)	.88**(*123*)	.83**(*24*)	-	.16(*44*)
48-59 Months (*N*)	.93**(*60*)	.18(*15*)	-	.38(*16*)
60-71 Months (*N*)	.95**(*27*)	.68*(*12*)	-	.80**(*15*)

*Note: * p < .05. ** p < .01. (2-tailed).*

Binary Logistic Regression Analysis of BADEC Versions for All Age Groups

The BADEC score, ADOS revised algorithm total score and the ADIR-T/ADI-R total score were used as individual predictors in the model to predict DSM-5 classification in participants (Table 16). The BADEC score was a significant predictor for all age groups except the 60-71 months old group. The ADOS revised algorithm total score was a significant predictor for only 24-35 and 36-47 months old group while the ADIR-T score was only significant for 36-47 months group.

Table 16

Binary Logistic Regression Analysis using BADEC, ADOS and ADI-R Score as

Individual Predictors

Age Group	Predictor	df, N	χ2	Odd Ratios	95% CI
12-23	**BADEC**	1, 114	56.48	2.23**	1.64, 3.03
Months	ADOS	1, 15	6.48	1.40	.97, 2.02
	ADIR-T	1, 17	7.01	1.76	.84, 3.70
24-35	**BADEC**	1, 133	107.98	2.63**	1.91, 3.61
Months	**ADOS**	1, 25	24.22	2.26*	1.19, 4.28
	ADIR-T	1, 36	16.36	2.38	.94, 6.03
36-47	**BADEC**	1, 123	57.55	2.13**	1.60, 2.82
Months	**ADOS**	1, 24	12.30	1.64*	1.06, 2.52
	ADIR-T	1, 44	14.67	1.35**	1.09, 1.66
48-59	**BADEC**	1,60	17.36	1.69**	1.21, 2.35
Months	ADOS	1, 15	12.66	1.99	.85, 4.68
	ADI-R	1, 16	1.32	1.05	.96, 1.15
60-71	BADEC	1, 27	12.85	2.78	.87, 8.82
Months	ADOS	1, 12	4.58	1.36	.97, 1.92
	ADI-R	1, 15	10.87	1.35	.97, 1.88

Note: $*p < .05$. $** p < .01$. Significant predictors in bold.

Discussion

Over the years, screening tools have been developed to identify children with ASD as early as possible. One of the available screening tools is the ADEC. Despite its robust psychometric properties, it is not known whether there are items more salient at various ages (i.e. critical items), and its usefulness with children older than 36 months old. Further, it is considered by some too time consuming for clinicians to use.

In this study, I had two aims. First, I examined the frequency and pattern of ASD behaviours that were observed in children with and without ASD based on the ADEC across the different age groups (12-23 months, 24-35 months, 36-47 months, 48-59 months and 60-71 months). Second, I reduced the ADEC into a brief version (BADEC) for each age group. The BADEC could then be used in primary care settings by busy health care professionals as rapid screeners or *red flags* to serve as guides for referral.

In the first part of the study, I found that children with ASD presented with a lower frequency of typical behaviours compared to children without ASD across all the age groups. Upon further visual inspection of the data, it appeared that there were certain key typical behaviours (i.e., response to name, gaze switch, eye contact, social smile and use of gestures) that children with ASD were less likely to demonstrate across most of the age groups. In the second part of the study, I examined whether I could develop a brief age-specific version of the ADEC (BADEC) for the different age groups using these key behaviours on the ADEC. The analyses supported the use of these key behaviours (i.e., ADEC items) to form one BADEC version for all age groups, albeit with different

cutoff scores. One version or form rather than different versions/forms for different age group would clearly be advantageous for clinicians.

The BADEC versions had acceptable internal consistency, correlated well with the full version, and mostly have sensitivity and specificity exceeding 80%, with the exception of specificity in the 48-59 months group (.53) and the 60-71 months group (.75). The minimum sensitivity and specificity for a screening tool should be between 70% and 80% and, in this case, the BADEC versions (especially for the younger age groups) may be deemed to be useful and effective (Glascoe, 2005). Diagnostic validity using the different BADEC versions was also demonstrated with significant group differences between participants with and without ASD.

I found a significant correlation between the BADEC versions and the ADOS revised algorithm total scores and the ADOS Toddler scores, except for the 48-59 months age group. However, correlations with the ADI-R Toddler and the ADI-R scores were only significant for the age groups of 24-35 months and 60-71 months respectively. The non-significant findings with the ADIR-T and ADI-R could be due to the relatively small sample size (N = 17) in the 12-23 months group while the ADI-R items for the older age groups or higher verbal ability tap on some behaviours (such as showing interest in other children, offering comfort, compulsions/rituals, etc.) that could not be observed and scored on the BADEC versions.

In addition, results from the logistic regression analyses indicated that the BADEC versions' total scores (with the exception of the 60-71 months group) were able to predict DSM-5 ASD classification just as well as the more time-intensive ADOS and ADI-R diagnostic tools, with odds ratio ranging from

1.69 to 2.63, while the ADOS revised algorithm score and the ADIR-T score recorded odds ratio ranging from 1.35 to 2.26. In addition, the ADOS score emerged as a significant predictor only for the age groups of 24-35 and 36-47 months while the ADIR-T score was only significant for 36-47 months group. However, these results need to be interpreted with some caution due to the relatively small sample size of participants who had the ADOS/ADIR-T/ADI-R data.

In this study, I noted that the PPV was relatively low (.64) for the BADEC (12 to 23 months) version compared to the other age group versions (ranging from .85 to .94). In this case, the BADEC (12 to 23 months) version tends to over-identify younger children as being at risk of having ASD, with 60% of false positive cases below 18 months old. Some researchers have described similar problems with ASD screening before age two and especially before 18 months (Barton et al., 2012) where there was a greater possibility of a higher false positive rate (and therefore low PPV) for the younger children (Chawarska et al., 2007; Pandey et al., 2008). This may be because some young children show early developmental variations which may resolve later, perhaps explaining why false positive rates may be higher for this age group (Swinkels et al., 2006). It is also possible that milder variants of ASD, and children with a higher level of cognitive development could be missed at a young age (Dietz, Swinkels, van Daalen, van Engeland, & Buitelaar, 2006). Therefore, it may be necessary to repeat the screening when the children are at 24 months old as recommended by the American Academy of Pediatrics (AAP, 2006; Johnson et al., 2007).

My finding of low PPV for this age group is also consistent with another screening tool similar to the ADEC, the Screening Tool for Autism in Two-Year-Olds (STAT) used for children under 24 months of age (Stone et al., 2008). The authors found that the PPV was .56. Similarly, a parent-report screening instrument (ESAT: Early Screening of Autistic Traits; Swinkels et al., 2006) to identify ASD at an earlier age of 14 months found that the PPV was .25. It is acknowledged that a high false positive rate may cause unnecessary concerns to parents. However, research suggests that even though most of the children who falsely screen positive for ASD at 18 months, are often at risk of other development disorders (Pandey et al., 2008; Pierce et al., 2011), the BADEC (12 to 23 months) version may still be useful in screening for the younger population who may benefit from some form of early intervention.

There are other brief versions of some of the ASD screening tools, such as the Autism Spectrum Quotient – Children Version (AQ-Child for 4-11 years old; Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008) and the Quantitative Checklist for Autism in Toddlers (Q-CHAT for 18-24 months old; Allison et al., 2008) that have been developed (Allison, Auyeung, & Baron-Cohen, 2012). However these brief versions are criticized due to being based on parental report, known to be less reliable than direct observation in toddlers (Barton et al., 2012; Wiggins et al., 2007). Further these tools are not suitable for children below 18 months of age. Crais et al. (2014) reported that there are few ASD specific tools available to screen infants below 18 months. Given parents are reporting signs emerging from as young as 12 months (Stone et al., 2004), the BADEC (12 to 23 months) version could fill the gap in this area by providing a brief and direct observation tool.

Limitations

One limitation of this study is the relatively small sample size of children with ASD ($N = 22$) as compared to children with non-ASD ($N = 92$) in the youngest age group (12-23 months old), and the relatively small sample size of children without ASD in the older age group (48-71 months old). This could be due to the nature of my screening study where it was difficult to recruit very young participants who eventually received a diagnosis of ASD after the screening process. This inherent difficulty in recruiting and diagnosing very young children was also noted in the validation study of the ADOS Toddler (Lord et al., 2012). I found that the small percentage (23%) of children diagnosed with ASD in the youngest age group was similar to the percentage (25%) found in the ADOS-T sample. I also had difficulty recruiting the older participants where parents (of older children) may not be keen to participate in the screening study. It could be because most of the developmental concerns (such as speech delay) may have already been identified or resolved in the older children and parents see no need for any developmental/ASD screening. In addition, not all participants had the ADOS and ADI-R data as these participants were recruited from the university autism centre and not via the screening study. Future research should replicate this study using a prospective design using larger sample of children with ASD in the youngest age group and a larger sample of children without ASD in the older age groups in order to evaluate the generalisability of the present findings.

Because the BADEC (specifically for the 12-23 and 24-35 months) and ADOS-T are similar in terms of clinician-administration and age group, future research could also examine the relationship of the BADEC versions and the

ADOS-T in terms of establishing concurrent validity. Similarities and differences between the BADEC (12-23 months and 24-35 months) versions and the ADOS-T would need to be explored with larger sample given the insufficient ADOS-T ($N = 20$) data in this present study.

Lastly, the findings in this study should be viewed as exploratory in nature, given that there were two issues to be considered. One issue was that examining a subset of scored items that were embedded in a larger administration might differ from the performance of those items in isolation. Another issue was that it was not clear whether the BADEC versions were meant for use as a Level 1 screening tool or as a Level 2 screening tool. If the BADEC versions were designated to be appropriate tools for paediatric providers to administer during well-child visits, then the BADEC must be validated in a Level 1 (low risk, universal) sample prior to its use.

Conclusion

This study represents the first step in understanding how ASD symptomology, as reflected on the ADEC test performance, changes across age in children with and without ASD. This understanding guided the development of different age versions of a brief ASD screening instrument designed to help health care professionals in the referral pathway for ASD. Clinicians and paediatricians may find the BADEC versions to be a quick and suitable screening tool to help them to identify young children presenting with possible ASD in their practice settings.

CHAPTER 5

General Discussion

Once considered to be a rare condition, ASD is now known to be one of the most common, and certainly one of the most debilitating, childhood disorders (Fombonne, 2008). This increase in prevalence may be due to improved methods of detection and to a shift away from understanding autism as a narrowly defined, categorical disorder to understanding it as a spectrum of conditions that affect individuals differently (Wing, 1996). Given early identification and intervention of ASD can dramatically improve the outcome of ASD (Dawson & Burner, 2011), it is of paramount importance to identify children with ASD as early as practical (Reichow, 2012).

Because there are no medical or genetic diagnostic tests for ASD, it may be difficult for medical professionals to identify children early. Unlike a medical condition where a blood, X-ray or CT scan test may provide positive results or markers for the disorder, there is no definite test for ASD. Instead, paediatricians and specialists have to rely on parental report, clinical judgment, and the ability to recognize the behavioural characteristics that define ASD. In general, best practice for a comprehensive assessment is to use multiple sources of information (i.e., interview, observation, rating scales) and involve multiple informants whenever possible (e.g., both parents, teachers; Kim & Lord, 2012a; Volkmar et al., 2014).

In this thesis, three studies were presented to investigate the psychometric properties of a relatively new observation screening measure, the Autism Detection in Early Childhood (ADEC; Young, 2007) in the early

identification of young children with possible ASD. To summarise, Study 1

provided a comprehensive psychometric validation of the ADEC as a screening

tool for ASD. In this study, the data showed that the ADEC is an effective

screening tool to identify children with ASD ranging from 12 to 36 months. The

ADEC has impressive psychometric properties, and can be administered easily

and quickly by persons with minimal training and experience with ASD. When

compared to the other Level 2 screening tools (e.g., the CARS and STAT), the

ADEC performed at a similar level (in terms of sensitivity and specificity) to

them. In my study, I found that the specificity of the ADEC was higher (than the

other Level 2 screening tools) when individuals with severe levels of intellectual

disability were excluded. The ADEC also offers the relative advantage of taking

lesser time to administer and to score, and being suitable for toddlers 24 months

and below. Clinicians and paediatricians may find the ADEC to be a suitable

screening tool to help them to identify young children presenting with possible

ASD in their practice settings.

Because there is currently little information available about the validity

of ASD screening tools in predicting long term outcomes such as diagnostic

classification, Study 2 compared the predictive validity data of the ADEC

against a well-established screening tool, the CARS (Schopler et al., 1998), in

relation to diagnostic classifications, symptom severity and functioning level at

2 and 6 years following initial assessment. Results indicated that both tools

performed similarly in their ability to predict with some accuracy long term

outcomes such as diagnostic status and overall adaptive functioning in our

validation sample. This study also extends our understanding of the

psychometric properties of both the ADEC and the CARS and proposes both the

ADEC and the CARS to be suitable ASD screening tools to predict long term

outcomes. No other studies have been conducted to examine the validity of

using the Level 2 ASD screening tools to predict long term outcomes in children

with ASD.

The examination of the psychometric properties of the ADEC in Studies

1 and 2 indicated the ADEC is a reliable and valid screening tool. Despite its

robust psychometric properties, it was not known whether there were items that

were more useful at various ages (i.e., critical items). Nor was it clear how

useful it would be with children older than 36 months. These factors, together

with the knowledge that it would be considered too time consuming for

clinicians to use, underpinned the direction of Study 3. In Study 3, I assessed the

ADEC's suitability as a screening tool for the older children because more than

half of children with developmental disabilities (including ASD) are not

identified until they enter school, around age 4 years (Sices et al., 2003).

Another focus of Study 3 was to ensure that the ADEC was maximally sensitive

to any variations in symptoms observed in children of different ages. There may

be certain behavioural markers of ASD that may be present in some younger

children which we may not see in older children with ASD, and vice versa.

Precise information on the age at which different behavioural features are

evident in children evaluated for ASD should highlight opportunities to improve

the early detection of ASD.

In the first part of Study 3, I examined the frequency and pattern of ASD

behaviours that were observed in children with and without ASD based on the

ADEC across the different age groups (12-23 months, 24-35 months, 36-47

months, 48-59 months and 60-71 months). Based on visual inspection, the

participants with ASD presented with a lower frequency of typical behaviours

compared to the participants without ASD across all the age groups. Further,

there were some key behaviours (e.g., response to name, gaze switch, and social

smile) that children with ASD were less likely to demonstrate across most of the

age groups. However, there were other behaviours (e.g., imitation, functional

play, pretend play, and delayed language) that seemed to improve and may no

longer be suitable behaviour markers as the child gets older.

In the second part of this study, given that I could identify behaviours

with better diagnostic salience at different ages, I examined whether I could

develop a brief age-specific ADEC (BADEC) version for identifying children

with possible ASD using those key behaviours identified. My analyses

supported the use of those critical items (e.g., response to name and gaze switch)

identified across most of the age groups to form one BADEC version for all age

groups, albeit with different cutoff scores. The brief versions for the different

age groups had acceptable internal consistency, correlated well with the full

version, and mostly have sensitivity and specificity exceeding 80%. The

BADEC versions' total scores (with the exception of the 60-71 months group)

were able to predict DSM-5 ASD classification just as well as the more time-

intensive ADOS and ADI-R diagnostic tools. The potential strength of this study

lies in reducing the ADEC into short forms that can not only be used to identify

children with a possible DSM-5 ASD diagnosis, but are also practical for busy

professionals to use in their settings. No other studies have been done to

examine whether other Level 2 screening tools can be shortened into brief

versions and to use for screening young children based on DSM-5 criteria.

However, it should be noted that this final study is exploratory in nature and the

findings should be replicated in further studies before the BADEC versions can

be used in clinical settings.

Future Directions

There is ongoing debate as to which screening methods (universal

screening or ASD-specific screening) and screening tools are most effective for

identifying children with early signs of ASD. It is also debatable whether

adopting a developmental surveillance approach is better than using screening

tools in identifying young children for possible ASD. I would like to argue that

neither one approach is inherently 'better'. There is evidence that the use of

screening instruments in developmental surveillance improves the efficiency of

an instrument (Glascoe, 1999). A very recent study by Davis, Clifton and

Papadopoulos (2015) investigated the use of the Social Attention and

Communication Study (SACS) and/or ADEC to facilitate the early diagnosis of

ASD. The results indicated that overall, when using a positive result in either the

SACS or the ADEC as a diagnostic test for ASD, there was a sensitivity of

95.5%, specificity of 75%, a positive predictive value of 84% and a negative

predictive value of 92.3%. This study used both the SACS and the ADEC,

where possible, in order to compare their effectiveness and the data suggested

both these tools are effective, easily administered, and there seems to be no

benefit in terms of reliability of one tool over the other. Given this finding, I

propose that the ADEC has the potential to be a suitable screening tool that can be incorporated in developmental surveillance.

While progress has been made in recent years, early screening for ASD remains far from perfect. It should be noted that screening results are sample specific (Charman & Gotham, 2013). For instance, the effectiveness of a screening tool could depend on factors such as the child's characteristics (e.g., clinical diagnosis, IQ, age), and family factors (e.g., parental education, parental awareness of autism). Added considerations include cost-effectiveness of the tool and the impact of misclassification (i.e. false positives and false negatives). While the development and validation for a *perfect* screening tool continues (especially in view of the new DSM-5 criteria), we need to use whatever knowledge and research findings we currently possess to make an informed decision of how best to identify children for ASD. The eventual and desired result should be that these children receive appropriate and early intervention.

Although I have demonstrated in this thesis that characteristics of ASD can be identified in children as early as 12 months of age, early identification may depend on the characteristics of the child and the family, as mentioned earlier. Previous research has suggested that it may be challenging to apply screening instruments developed in Western countries into non-Western cultures (Wallis & Pinto-Martin, 2008). Some possible differences in features of ASD between Western and Eastern cultures have been reported in terms of eye contact and early language development (Bernier, Mao, & Yen, 2010; Daley & Sigman, 2002). For instance, in Asian cultures, looking directly into another (adult) person's eye may be considered as rude, threatening, or disrespectful.

However, avoidance of eye contact is one such core social communicative deficit that has been well reported in Western studies. In another example, there are some parents and grandparents of children with ASD in mainland China who consider boys speaking late to be a good sign for their future development (Sun et al., 2013).

For these various reasons, it is important to gather data on the clinical usefulness and validity of promising tools in culturally and linguistically diverse populations. In this thesis, the ethnic background of the sample was predominantly Caucasian (at least 90%), which may limit the generalizability of the findings to non-Caucasian population. Though the ADEC has been translated by the bilingual research team into Spanish and researched by a bilingual psychologist in Mexico (Hedley et al., 2010), and results have been encouraging, future research could also examine the use of the ADEC in other developing and non-English-speaking Asian cultures such as in China and India. For instance, researchers in Taiwan (Chiang et al., 2013) have modified the Screening Tool for Autism in Two-Year-Olds (STAT) into a Taiwanese version called T-STAT in Chinese, and results were promising as a screening tool for ASD for children aged two to three years old.

Future research could also look into developing and validating a brief parent version of the ADEC using those key behaviours identified in Study 3. Best practice in ASD assessment proposes using data from multiple sources (i.e., clinicians, caregivers, and teachers) to enhance accuracy for the diagnosis of ASD (Kim & Lord, 2012a). Since the psychometric properties of the ADEC as a clinician-administered interactive screening tool has been established, it

would be useful to have a parent-report version to complement the use of the ADEC. It has been suggested that an ideal screening tool should have these features besides excellent psychometric properties: quick, with a possible maximum of five questions, electronically available, easy to score, one-page including scoring, at a relatively low reading level, and culturally sensitive (Crais et al., 2014). With the advent of technology, it would now be possible to develop an electronic version or application software of a brief parent version of the ADEC that clinicians and medical professionals can use and integrate into their practice.

Conclusion

Over the last two decades, prevalence of ASD has been rising steeply. Clinicians and paediatricians are likely to see an increase in children presenting with ASD and will need appropriate tools and training to identify them. Many screening instruments for ASD have been developed, although few are well-evaluated.

The studies in this thesis represent the first step in understanding the psychometric properties and usefulness of an ASD-specific screening tool, the ADEC, in the early identification of young children presenting with possible ASD. The data from this thesis support the use of the ADEC to be a quick and suitable screening tool by clinicians and paediatricians to help them to identify these children in their practice settings. With a clearer understanding of the ADEC's psychometric properties and usefulness in early detection and diagnosis, it can be then used in a proper context (such as in early intervention

programs, or evaluation clinics serving children with a variety of developmental

problems) to identify young toddlers who will benefit from early intervention.

References

Abrahams, B., & Geschwind, D. (2008). Advances in autism genetics: On the threshold of a new neurobiology. *Nature Reviews. Genetics, 9*, 341-355. doi:10.1038/nrg2346

Allison, S., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for Autism screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist in 1,000 cases and 3,000 controls. *Journal of The American Academy Of Child & Adolescent Psychiatry, 51*, 202-212. doi:10.1016/j.jaac.2011.11.003

Allison, S., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G., & Brayne, C. (2008). The Q-CHAT (Quantitative Checklist for Autism in Toddlers): A normally distributed quantitative measure of autistic traits at 18-24 months of age: Preliminary report. *Journal of Autism and Developmental Disorders, 38*, 1414-1425. doi:10.1007/s10803-007-0509-7

American Academy of Pediatrics. (2003). Periodic Survey of Fellows #53: Identification of children <36 months at risk for developmental problems and referral to early identification programs. Retrieved January 6, 2014 from http://www.aap.org/research/periodicsurvey/ps53exs.htm.

American Academy of Pediatrics. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, *118*, 405–420. doi:10.1542/peds.2006-1231

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders: Text Revision (DSM-IV-TR)* (4th ed.). Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: Fifth Edition (DSM-5)*. Arlington, VA: Author.

Anderson, G. M. (2014). Autism Biomarkers: Challenges, Pitfalls and Possibilities. *Journal of Autism and Developmental Disorders*, 1-11. doi:10.1007/s10803-014-2225-4

Australian Institute of Health and Welfare. (2009). *The national health performance framework*. Retrieved from http://meteor.aihw.gov.au/content/index.phtml/itemId/435314NHpFAIo HaWM]

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The Autism Spectrum Quotient: Children's Version (AQ-Child). *Journal of Autism and Developmental Disorders, 38*, 1230-1240. doi:10.1007/s10803-007-0504-z

Aylward, G. P. (1997). Conceptual issues in developmental screening and assessment. *Journal of Developmental & Behavioral Pediatrics*, *18*, 340-349.

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 and Adolescent Psychiatry, 39,* 694–702. doi:10.1097/00004583-200006000-00007

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*, 468-475. doi:10.1136/adc.84.6.468

Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory, motor and social behaviours at 9–12 months of age. *Journal of Autism and Developmental Disorders, 29*, 213–224. doi:10.1023/A:1023080005650

Barbaro, J., & Dissanayake, C. (2010). Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: the social attention and communication study. *Journal of Developmental & Behavioral Pediatrics*, *31*, 376-385. doi:10.1097/DBP.0b013e3181df7f3c

Barbaro, J., Ridgway, L., & Dissanayake, C. (2011). Developmental surveillance of infants and toddlers by maternal and child health nurses in an Australian community-based setting: promoting the early identification of autism spectrum disorders. *Journal of Pediatric Nursing*, *26*, 334-347. doi:10.1016/j.pedn.2010.04.007

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*, 1165–1174. doi:10.1007/s10803-011-1343-5

Bayley, N. (2005). *Bayley Scales of Infant and Toddler Development (3rd ed.) Screening Test*. San Antonio, TX: Pearson.

Bernier, R., Mao, A., & Yen, J. (2010). Psychopathology, families, and culture: Autism. *Child and Adolescent Psychiatric Clinics of North America, 19*, 855–867. doi:10.1016/j.chc.2010.07.005

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *The British Journal of Psychiatry, 175*, 444-451. doi:10.1192/bjp.175.5.444

Bishop, S. L., Luyster, R., Richler, J., & Lord, C. (2008). Diagnostic assessment. In: K. Chawarska, A. Klin, & F.R. Volkmar. *Autism Spectrum Disorder in infants and toddlers: diagnosis, assessment, and treatment.* New York: The Guilford Press.

Bruni, T. P. (2014). Test Review: Social Responsiveness Scale–Second Edition (SRS-2). *Journal of Psychoeducational Assessment*, *32*, 365-369. doi:10.1177/0734282913517525

Camp, B. W. (2006). What the clinician really needs to know: Questioning the clinical usefulness of sensitivity and specificity in studies of screening tests. *Journal of Developmental & Behavioral Pediatrics*, *27*, 226-230.

Centers for Disease Control and Prevention (2002). Prevalence of autism spectrum disorders—autism and developmental disabilities monitoring network, United States. *Morbidity and Mortality Weekly Report Surveillance Summary, 58*, 1–20.

Centers for Disease Control and Prevention. (2012). Prevalence of Autism Spectrum Disorders: Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. *Morbidity and Mortality Weekly Report, 61*, 1-18.

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism

spectrum disorder in 2- and 3- year-old children. *Journal of Child*

Psychology and Psychiatry, 43, 289–305. doi:10.1111/1469-7610.00022

Charman, T., & Baron-Cohen, S. (2006). Screening for autism spectrum

disorders in populations: Progress, challenges, and questions for future

research and practice. In T. Charman & W. Stone (Eds.), *Social and*

communication development in autism spectrum disorders (pp. 63–82).

New York: Guilford Press.

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and

diagnostic instruments for autism spectrum disorders – lessons from

research and practice. *Child and Adolescent Mental Health, 18*, 52–64.

doi:10.1111/j.1475-3588.2012.00664.x

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G.

(2011). IQ in children with autism spectrum disorders: Data from the

Special Needs and Autism Project (SNAP). *Psychological Medicine*, *41*,

619–627. doi:10.1017/S0033291710000991

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J. A., & Baird, G.

(2005). Outcome at 7 years of children diagnosed with autism at age 2:

Predictive validity of assessments conducted at 2 and 3 years of age and

pattern of symptom change over time. *Journal of Child Psychology and*

Psychiatry, 46, 500–513. doi:10.1111/j.1469-7610.2004.00377.x

Chawarska, K., & Bearss, K. (2008). Assessment of cognitive and adaptive

 skills. In K. Chawarska, A. Klin, & F. R. Volkmar (Eds.), *Autism*

 spectrum disorders in infants and toddlers: Diagnosis, assessment, and

 treatment (pp. 50–75). New York: Guilford Press.

Chawarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2009). A

 prospective study of toddlers with ASD: Short-term diagnostic and

 cognitive outcomes. *Journal of Child Psychology and Psychiatry, 50*,

 1235–1245. doi:10.1111/j.1469-7610.2009.02101.x

Chawarska, K., Klin, A., Paul, K., & Volkmar, F. (2007). Autism spectrum

 disorder in the second year: Stability and change in syndrome

 expression. *Journal of Child Psychology and Psychiatry, 48*, 128-138.

 doi:10.1111/j.1469-7610.2006.01685.x

Chawarska, K., & Volkmar, F. (2005). Autism in infancy and early childhood.

 In: F. Volkmar, A. Klin, R. Paul, & D.J. Cohen (Eds.), *Handbook of*

 autism and pervasive developmental disorders (3rd ed., Vol. 1, pp. 223-

 246). New York: Wiley.

Chiang, C. H., Wu, C. C., Hou, Y. M., Chu, C. L., Liu, J. H., & Soong, W. T.

 (2013). Development of T-STAT for early autism screening. *Journal of*

 Autism and Developmental Disorders, *43*, 1028-1037.

 doi:10.1007/s10803-012-1643-4

Chlebowski, C., Green, J. A., Barton, M. L., & Fein, D. (2010). Using the

 Childhood Autism Rating Scale to diagnose Autism Spectrum Disorders.

 Journal of Autism and Developmental Disorders, 40, 787-799.

 doi:10.1007/s10803-009-0926-x

Clark, A., & Harrington, R. (1999). On diagnosing rare disorders rarely:

    Appropriate use of screening instruments. *Journal of Child Psychology*

    and Psychiatry, *40*, 287-290. doi:10.1111/1469-7610.00442

Clifford, S., Young, R., & Williamson, P. (2007). Assessing the early

    characteristics of Autistic Disorder using video analysis. *Journal of*

    Autism and Developmental Disorders, 37, 301-313.

    doi:10.1007/s10803-006-0160-8

Constantino, J. N., & Gruber, C. P. (2005). *Social Responsiveness Scale (SRS).*

    Los Angeles, LA: Western Psychological Services.

Constantino, J. N., & Gruber, C. P. (2012). *Social Responsiveness Scale–Second*

    Edition (SRS-2). Torrance, CA: Western Psychological Services.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., &

    Przybeck, T. (2004). The factor structure of autistic traits. *Journal of*

    Child Psychology and Psychiatry, 45, 719–726. doi:10.1111/j.1469-

    7610.2004.00266.x

Constantino, J. N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna,

    K., . . . Todd, R. D. (2006). Autistic social impairment in the siblings of

    children with pervasive developmental disorders. *American Journal of*

    Psychiatry, 163, 294–296. doi:10.1176/appi.ajp.163.2.294

Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., . .

    . Nightingale, N. (1999). Autism spectrum disorders at 20 and 42 months

    of age: Stability of clinical and ADI-R diagnosis. *Journal of Child*

    Psychology and Psychiatry, 40, 719–732. doi:10.1111/1469-7610.00488

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, 2311-2328. doi:10.1007/s10803-014-2101-2

Daley, T. C., & Sigman, M. D. (2002). Diagnostic conceptualization of autism among Indian psychiatrists, psychologists, and paediatricians. *Journal of Autism and Developmental Disorders, 32*, 13–23. doi:10.1023/A:1017947922349

Dawson, G., & Burner, K. (2011). Behavioral interventions in children and adolescents with autism spectrum disorder: a review of recent findings. *Current Opinion in Pediatrics*, *23*, 616-620. doi:10.1097/MOP.0b013e32834cf082

Davis, T., Clifton, D., & Papadopoulos, C. (2015). Identifying autism early: The Toddlers at Risk of Autism Clinic model. *Journal of Paediatrics and Child Health*. Advance online publication. doi:10.1111/jpc.12832

Dietz, C., Swinkels, S. H. N., van Daalen, E., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum disorder in children aged 14–15 months. II: Population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. *Journal of Autism and Developmental Disorders, 36*, 713–722. doi:10.1007/s10803-006-0114-1

De Bildt, A., Sytema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., &

    Minderaa, R. (2004). Interrelationship between Autism Diagnostic

    Observation Schedule-Generic (ADOS-G), Autism Diagnostic

    Interview-Revised (ADI-R), and the Diagnostic and Statistical Manual of

    Mental Disorders (DSM-IV-TR) classification in children and

    adolescents with Mental Retardation. *Journal of Autism and*

    Developmental Disorders, 34, 129–137.

    doi:10.1023/B:JADD.0000022604.22374.5f

DiLalla, D. L., & Rogers, S. J. (1994). Domains of the Childhood Autism

    Rating Scale: Relevance for diagnosis and treatment. *Journal of Autism*

    and Developmental Disorders, 24, 115-128. doi:10.1007/BF02172092

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 & Behavioral Pediatrics*,

    27(2, Suppl.), S88–S94. doi:10.1097/00004703-200604002-00006

Douglas, K. S., Guy, L. S., Reeves, K. A., & Weir, J. (2008). *HCR-20 violence*

    risk assessment scheme: Overview and annotated bibliography.

    Retrieved from http://escholarship.umassmed.edu/psych_cmhsr/335/

Eaves, L., & Ho, H. (2004). Brief report: Stability and change in cognitive and

    behavioural characteristics of autism through childhood. *Journal of*

    Autism and Developmental Disorders 26, 557–569.

    doi:10.1007/BF02172276

Eaves, L., Wingert, H., Ho, H.., & Mickelson, E. (2006). Screening for autism

    spectrum disorders with the social communication questionnaire.

    Developmental and Behavioral Pediatrics, 27, 95-103.

    doi:10.1097/00004703-200604002-00007

Eaves, R., & Milner, B. (1993). The criterion-related validity of the childhood

    autism rating scale and the autism behavior checklist. *Journal of*

    Abnormal Child Psychology, 21, 481–491. doi:10.1007/BF00916315

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. *Autism Research, 5*, 160–179.

    doi:10.1002/aur.239

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999).

    Evaluating the use of exploratory factor analysis in psychological

    research. *Psychological Methods*, *4*(3), 272.

Feldman, M. A., Ward, R. A., Savona, D., Regehr, K., Parker, K., Hudson, M., .

    . . Holden, J. A. (2012). Development and initial validation of a parent

    report measure of the behavioral development of infants at risk for

    Autism Spectrum Disorders. *Journal of Autism and Developmental*

    Disorders, 42, 13–22. doi: 10.1007/s10803-011-1208-y

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, Jr., 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. (1999). The epidemiology of autism: A review. *Psychological Medicine*, *29*, 769–786.

Fombonne, E. (2005). The changing epidemiology of Autism. *Journal of Applied Research in Intellectual Disabilities, 18*, 281-294. doi:10.1111/j.1468-3148.2005.00266.x

Fombonne, E. (2008). Is autism getting commoner? *British Journal of Psychiatry, 193*, 59. doi:10.1192/bjp.193.1.59

Fombonne, E. (2009). Epidemiology of Pervasive Developmental Disorders. *Pediatric Research, 65*, 591-598. doi:10.1203/PDR.0b013e31819e7203

Fombonne, E., Quirke, S., & Hagen, A. (2011). Epidemiology of pervasive developmental disorders. In D. G. Amaral, G. Dawson, & D. H. Geschwind (Eds.), *Autism spectrum disorders* (pp. 90–111). New York: Oxford University Press.

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders, 38*, 474–480. doi:10.1007/s10803-007-0415-z

Frazier, T., Youngstrom, E., 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 and Adolescent Psychiatry, 51*, 28–40. doi:10.1016/j.jaac.2011.09.021

García-Primo, P., Hellendoorn, A., Charman, T., Roeyers, H., Dereu, M., Roge, B., . . . & Canal-Bedia, R. (2014). Screening for autism spectrum disorders: state of the art in Europe. *European Child & Adolescent Psychiatry*, *23*, 1005-1021. doi:10.1007/s00787-014-0555-6

Geier, D. A., Kern, J. K., & Giier, M. R. (2013). A comparison of the Autism Treatment Evaluation Checklist (ATEC) and the Childhood Autism Rating Scale (CARS) for the quantitative evaluation of Autism. *Journal of Mental Health Research in Intellectual Disabilities, 6*, 255-267. doi:10.1080/19315864.2012.681340

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*, 1750–1756. doi:10.1007/s10803-012-1560-6

Gillis, J. (2009). Screening practices of family physicians and pediatricians in 2 southern states. *Infants & Young Children*, *22*(4), 321–331. doi:10.1097/IYC.0b013e3181bc4e21

Glascoe, F. P. (1999). The value of parents' concerns to detect and address developmental and behavioural problems. *Journal of Paediatrics and Child Health*, *35*, 1-8. doi:10.1046/j.1440-1754.1999.00342.x

Glascoe, F.P. (2005) Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities, 11*, 173–179. doi:10.1002/mrdd.20068

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 693–705. doi:10.1007/s10803-008-0674-3

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule (ADOS): Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37*, 613–627. doi:10.1007/s10803-006-0280-1

Gray, K. M., Tonge, B. J. & Sweeney, D. J. (2008). Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: Evaluating diagnostic validity. *Journal of Autism and Developmental Disorders, 38*, 657–667. doi:10.1007/s10803-007-0432-y

Gura, G. F., Champagne, M. T., & Blood-Siegfried, J. E. (2011). Autism spectrum disorder screening in primary care. *Journal of Developmental & Behavioral Pediatrics*, *32*, 48-51.

Harrison, P. L. & Oakland, T. (2003). *Adaptive Behaviour Assessment System* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Hedley, D., Young, R., Juarez-Gallegos, M., & Marcin-Salazar, C. (2010). Cross-cultural evaluation of the Autism Detection in Early Childhood (ADEC) in Mexico. *Autism, 14*, 93-112. doi:10.1177/1362361309347676

Helt, M., Kelley, E., Kinsbourne, M., Pandey, J., Boorstein, H., Herbert, M., & Fein, D. (2008). Can children with autism recover? If so, how? *Neuropsychology Review, 18*, 339–366. doi:10.1007/s11065-008-9075-9

Honigfeld, L., & McKay, K. (2006). Barriers to enhancing practice based

    developmental services. *Journal of Developmental and Behavioral*

    Pediatrics, 27(1 Suppl), S30–33; discussion S34–37, S50–32.

Huerta, M., Bishop, S., Duncan, A., Hus, V., & Lord, C. (2012). Application of

    DSM-5 criteria for Autism Spectrum Disorder to three samples of

    children with DSM-IV diagnoses of Pervasive Developmental Disorders.

    The American Journal of Psychiatry, 169, 1056-1064.

    doi:10.1176/appi.ajp.2012.12020276

Johnson, C. P., Meyers, S. M., & the Council on Children with Disabilities

    (2007). Identification and evaluation of children with autism spectrum

    disorders. *American Academy of Pediatrics, 120*, 1183-1215.

    doi:10.1542/peds.2007-2361

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-

    month-old infants later diagnosed with autism. *Nature, 504*, 427–431.

    doi:10.1038/nature12715

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and

    social-communicative functioning in children with autism spectrum

    disorder. *Journal of Child Psychology and Psychiatry, and Allied*

    Disciplines, 43, 807. doi:10.1111/1469-7610.00092

Kazdin, A. E. (1994). Evaluation in clinical practice: Clinically sensitive and

    systematic methods of treatment delivery. *Behavior Therapy*, *24*, 11-45.

    doi:10.1016/S0005-7894(05)80255-2

Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., & Hellemann, G. (2012). Longitudinal follow-up of children with Autism receiving targeted interventions on joint attention and play. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 487-495. doi:10.1016/j.jaac.2012.02.019

Kim, S. H., & Lord, C. (2012a). Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Child Psychology and Psychiatry*, *53*, 143-151. doi: 10.1111/j.1469-7610.2011.02458.x

Kim, S. H., & Lord, C. (2012b). New Autism Diagnostic Interview-Revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Autism and Developmental Disorders, 42*, 82-93. doi:10.1007/s10803-011-1213-1

Kim, S. H., Thurm, A., Shumway, S., & Lord, C. (2012). Multisite study of new Autism Diagnostic Interview-Revised (ADI-R) algorithms for toddlers and young preschoolers. *Journal of Autism and Developmental Disorders*, 1-12. doi:10.1007/s10803-012-1696-4

King, M., & Bearman, P. (2009). Diagnostic change and the increased prevalence of autism. *International Journal of Epidemiology, 38*, 1224–1234. doi:10.1093/ije/dyp261

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

and Developmental Disorders, 30, 163–167.

doi:10.1023/A:1005415823867

Kong, S. W., Collins, C. D., Shimizu-Motohashi, Y., Holm, I. A., Campbell, M.

G., Lee, I. H., . . . & Kohane, I. S. (2012). Characteristics and predictive

value of blood transcriptome signature in males with autism spectrum

disorders. *PLoS One*, *7*(12), e49475. doi:10.1371/journal.pone.0049475

Kulage, K. M., Smaldone, A. M., & Cohn, E. G. (2014). How will DSM-5 affect

autism diagnosis? A systematic literature review and meta-analysis.

Journal of Autism and Developmental Disorders, *44*, 1918-1932.

doi:10.1007/s10803-014-2065-2

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement

for categorical data. *Biometrics, 33,* 159–174. doi:10.2307/2529310

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The Autism Diagnostic Interview

– Revised (ADI-R).* Los Angeles, CA: Western Psychological Services.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism.

Journal of Child and Adolescent Psychiatry, 36, 1365– 1382.

Lord, C., Luyster, R., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic

Observation Schedule, Second Edition (ADOS-2) Manual (Part II):

Toddler Module*. Torrance, CA: Western Psychological Services.

Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D., M., . . . Risi, S.

(2012). A multisite study of the clinical diagnosis of different Autism

Spectrum Disorders. *Archives of General Psychiatry, 69*, 306-313.

doi:10.1001/archgenpsychiatry.2011.148

Lord, C., Risi, S., DiLavore, P., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from two to nine. *Archives of General Psychiatry, 63*, 694–701. doi:10.1001/archpsyc.63.6.694

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205–223. doi: 10.1023/A:1005592401947

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1-4*. Torrance, CA: Western Psychological Services.

Lord, C., Shulman, C., & DiLavore, P. (2004). Regression and word loss in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry, 45*, 936–955. doi:10.1111/j.1469-7610.2004.t01-1-00287.x

Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal, 14*, 234–252. doi:10.1002/1097-0355(199323)14:3<234::AID-IMHJ2280140308>3.0.CO;2-F

Luiselli, J. K., Campbell, S., Cannon, B., DiPietro, E., Ellis, J. T., Taras, M., & Lifter, K. (2001). Assessment instruments used in the education and treatment of persons with autism: Brief report of a survey of national service centers. *Research in Developmental Disabilities*, *22*, 389-398. doi:10.1016/S0891-4222(01)00079-8

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., . . .
Lord, C. (2009). The Autism Diagnostic Observation Schedule—Toddler
Module: A new module of a standardized diagnostic measure for Autism
Spectrum Disorders. *Journal of Autism and Developmental Disorders,
39,* 1305-1320. doi:10.1007/s10803-009-0746-z

Maljaars, J., Noens, I., Scholte, E., & van Berckelaer-Onnes, I. (2012).
Evaluation of the criterion and convergent validity of the Diagnostic
Interview for Social and Communication Disorders in young and low-
functioning children. *Autism, 16*, 487-497.
doi:10.1177/1362361311402857

Matson, J. L., Belva, B. C., Horovitz, M., & Bamburg, J. (2012). Comparing
symptoms of autism spectrum disorders in a developmentally disabled
adult population using the current DSM-IV-TR diagnostic criteria and
the proposed DSM-5 diagnostic criteria. *Journal of Developmental and
Physical Disabilites*. doi:10.1007/s10882-012-9278-0.

Matson, J. L., Boisjoli, J. A., & Wilkins, J. (2007). *The baby and infant screen
for children with autism traits (BISCUIT).* Baton Rouge, LA: Disability
Consultants, LLC.

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

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its
relationship to autism spectrum disorders. *Research in Developmental
Disabilities, 30*, 1107–1114. doi:10.1016/j.ridd.2009.06.003

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,* 1139–1146. doi:10.1007/s10803-010-0973-3

Matson, J. L., Wilkins, J., Sevin, J. A., Knight, C., Boisjoli, J. A., & Sharp, B. (2009a). 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*, 336–344. doi:10.1016/j.rasd.2008.08.001

Matson, J. L., Wilkins, J., Sharp, B., Knight, C., Sevin, J. A., & Boisjoli, J. A. (2009b). 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*, 924–930. doi: 10.1016/j.rasd.2009.04.001

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*, 368-383. doi:10.1016/j.jaac.2012.01.007

Menard, S. (1995). *Applied logistic regression analysis*. Sage university paper series on quantitative application in the social sciences, 07-106. Thousand Oaks, CA: Sage.

Miller, J. S., Gabrielsen, T., Villalobos, M., Alleman, R., Wahmhoff, N., Carbone, P. S., & Segura, B. (2011). The each child study: Systematic screening for autism spectrum disorders in a pediatric setting. *Pediatrics*, *127*, 866-871. doi:10.1542/peds.2010-0136

Mordre, M., Groholt, B., Knudsen, A., Sponheim, E., Mykletun, A., & Myhre, A. (2012). Is long term prognosis for pervasive developmental disorder not otherwise specified different from prognosis for autistic disorder? Findings from a 30 year follow up study. *Journal of Autism and Developmental Disorders, 42*, 920-928. doi:10.1007/s10803-011-1319-5

Moore, V., & Goodson, S. (2003). How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. *Autism, 7,* 47–63. doi:10.1177/1362361303007001018

Mullen, E. M. (1995). *Mullen Scales of Early Learning: AGS edition*. Circle Pines, MN: American Guidance Service.

Munson, J., Dawson, G., Sterling, L., Beauchaine, T., Zhou, A., Koehler, E., . . . Conners, F. (2008) Evidence for latent classes of IQ in young children with Autism Spectrum Disorder. *American Journal on Mental Retardation, 113*, 439-452. doi:10.1352/2008.113:439–452

Myers, R. (1990). *Classical and modern regression with applications* (2nd ed.). Boston, MA: Duxbury.

Nah, Y. H., Young, R. L., & Brewer, N. (2014a). Using the Autism Detection in Early Childhood (ADEC) and Childhood Autism Rating Scales (CARS) to predict long term outcomes in children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 44*, 2301-2310. doi:10.1007/s10803-014-2102-1

Nah, Y. H., Young, R. L., Brewer, N., & Berlingeri, G. (2014b). Autism

Detection in Early Childhood (ADEC): Reliability and validity data for a

level 2 screening tool for Autistic Disorder. *Psychological Assessment,*

26, 215-226. doi:10.1037/a0034472

Nordin, V., Gillberg, C., & Nyde´n, A. (1998). The Swedish version of the

childhood autism rating scale in a clinical setting. *Journal of Autism and*

Developmental Disorders, 28, 69–75. doi:10.1023/A:1026067104198

Norris, M., & Lecavalier, L. (2010). Screening accuracy of Level 2 Autism

Spectrum Disorder rating scales: A review of selected instruments.

Autism, *14*, 263-284. doi:10.1177/1362361309348071

Oosterling, I., Roos, S., de Bildt, A., Rommelse, N., de Jonge, M., Visser, J., . . .

Buitelaar, J. (2010). Improved diagnostic validity of the ADOS revised

algorithms: A replication study in an independent sample. *Journal of*

Autism and Developmental Disorders, 40, 689–703.

doi:10.1007/s10803-009-0915-0

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based

assessment of Autism Spectrum Disorders in children and adolescents.

Journal of Clinical Child and Adolescent Psychology, 34, 523-540.

Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A., Babitz, T., . . . &

Fein, D. (2008). Screening for autism in older and younger toddlers with

the Modified Checklist for Autism in Toddlers. *Autism*, *12*, 513-535.

doi:10.1177/1362361308094503

Parner, E. T., Thorsen, P., Dixon, G., de Klerk, N., Leonard, H., Nassar, N., . . .

> Glasson, E. J. (2011). A comparison of autism prevalence trends in

> Denmark and Western Australia. *Journal of Autism and Developmental*

> *Disorders, 41*, 1601–1608. doi:10.1007/s10803-011-1186-0

Perry, A., Condillac, R., Freeman, N., Dunn-Geier, J., & Belair, J. (2005).

> Multi-site study of the Childhood Autism Rating Scales (CARS) in five

> clinical groups of young children. *Journal of Autism and Developmental*

> *Disorders, 35*, 625–634. doi:10.1007/s10803-005-0006-9

Perry, A., & Freeman, N. (1996). *Psychometric properties of the Childhood*

> *Autism Rating Scale (CARS).* Poster presented at the Ontario Association

> on Developmental Disabilities Research Special Interest Group Research

> Day, Toronto, ON.

Pierce, K., Carter, C., Weinfeld, M., Desmond, J., Hazin, R., Bjork, R., &

> Gallagher, N. (2011). Detecting, studying, and treating autism early: the

> one-year well-baby check-up approach. *The Journal of Pediatrics*, *159*,

> 458-465. doi:10.1016/j.jpeds.2011.02.036

Pilowsky, T., Yirmiya, N., Shulman, C., & Dover, R. (1998). The Autism

> Diagnostic Interview-Revised and the Childhood Autism Rating Scale:

> Differences between diagnostic systems and comparison between

> genders. *Journal of Autism and Developmental Disorders, 28*, 143-151.

> doi:10.1023/A:1026092632466

Reichow, B. (2012). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*, 512-520. doi:10.1007/s10803-011-1218-9

Rellini, E., Tortolani, D., Trillo,S., Carbone, S., & Montecchi, F. (2004). Childhood Autism Rating Scale (CARS) and Autism Behavior Checklist (ABC) correspondence and conflicts with DSM-IV criteria in diagnosis of Autism. *Journal of Autism and Developmental Disorders, 34*, 703-708. doi:10.1007/s10803-004-5290-2

Riegelman, R. K. & Hirsch, R. P. (1989). *Studying a study and testing a test: How to read the medical literature*. Boston: Little, Brown & Co.

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 and Behavioral Pediatrics, 27*, 111–119. doi:10.1097/00004703-200604002-00009

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,* 131–144. doi:10.1023/A:1010738829569

Rust, J. O., & Wallace, M. A. (2004). Adaptive Behavior Assessment System–Second edition. *Journal of Psychoeducational Assessment, 22*, 367-373.

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

Schalock, R. L., Buntinx, W. H. E., Borthwick- Duffy, S., Luckasson, R., Snell, M. E., Tasse´, M. J., & Wehmeyer, M. L. (2007). *User's guide: Mental retardation: Definition, classification, and systems of supports. Applications for clinicians, educators, disability program managers, and policy makers*. Washington, DC: American Association on Intellectual and Developmental Disabilities.

Schopler, E., Bourgondien, M. E. V., Wellman, G. J., & Love, S. R. (2010). *The Childhood Autism Rating Scale Second Edition (CARS2).* Los Angeles: Western Psychological Services.

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

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., . . . Cuniff, C. (2009). Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*, 474-483. doi:10.1097/CHI.0b013e31819b3848

Sices, L., Feudtner, C., McLaughlin, J., Drotar, D., & Williams, M. (2003). How do primary care physicians identify young children with developmental delays? A national survey. *Journal of Developmental and Behavioral Pediatrics, 24*(6), 409–417.

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

Sivberg, B. (2003). Parents' detection of early signs in their children having an
    autism spectrum disorder. *Journal of Pediatric Nursing, 18*, 433-439.
    doi:10.1016/S0882-5963(03)00139-8

Snow, A., Lecavalier, L., & Houts, C. (2009). Structure of the autism diagnostic
    interview-revised: Diagnostic and phenotypic implications. *Journal of
    Child Psychology and Psychiatry, 50*, 734–742. doi:10.1111/j.1469-
    7610.2008.02018.x

Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland Adaptive Behavior
    Scales* (Survey Form). Circle Pines, MN: American Guidance Service.

Sparrow, S., Cicchetti, D., & Balla, D. (2005). *Vineland Adaptive Behavior
    Scales* (2nd ed.). Minneapolis, MN: Pearson Assessment.

Steiner, A. M., Goldsmith, T. R., Snow, A. V., Chawarska, K. (2012).
    Practitioner's guide to assessment of Autism Spectrum Disorders in
    infants and toddlers. *Journal of Autism and Developmental Disorders,
    42*, 1183–1196. doi:10.1007/s10803-011-1376-9

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Screening tool for autism
    in two-year-olds (STAT): Development and preliminary data. *Journal of
    Autism and Developmental Disorders, 30*, 607–612.
    doi:10.1023/A:1005647629002

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*, 691– 701.
    doi:10.1007/s10803-004-5289-8

Stone, W. L., & Hogan, K. (1993). A structured parent interview for identifying young children with autism. *Journal of Autism and Developmental Disorders, 23*, 639–652. doi:10.1007/BF01046106

Stone, W. L., Lee, E., Ashford, L., Brissie, J., Hepburn, S., Coonrod, E., & Weiss, B. (1999). Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 219–226. doi:10.1111/1469-7610.00435

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*, 557-573. doi:10.1177/1362361308096403

Stone, W. L., & Yoder, P. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism, 5*, 341-361. doi:10.1177/1362361301005004002

Sun, X., Allison, C., Auyeung, B., Matthews, F. E., Baron-Cohen, S., & Brayne, C. (2013). Service provision for autism in mainland China: A service providers' perspective. *Research in Developmental Disabilities, 34*, 440–451. doi:10.1016/j.ridd.2012.08.010

Swinkels, S. H. N., Dietz, C., van Daalen, E., Kerkhof, I. H. G. M., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum in children aged 14 to 15 months. I: The development of the early screening of autistic traits questionnaire (ESAT). *Journal of Autism and Developmental Disorders, 36*, 723–732. doi:10.1007/s10803-006-0115-0

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.).

    Boston: Allyn & Bacon.

Tachimori, H., Osada, H., & Kurita, H. (2003). Childhood Autism Rating

    Scale—Tokyo Version for screening pervasive developmental disorders.

    Psychiatry and Clinical Neurosciences, 57, 113-118.

    doi:10.1046/j.1440-1819.2003.01087.x

Thurm, A., Lord, C., Lee, L., & Newschaffer, C. (2007). Predictors of language

    acquisition in preschool children with autism spectrum disorders.

    Journal of Autism and Developmental Disorders, 37, 1721-1734.

    doi:10.1007/s10803-006-0300-1

Van Daalen, E., Kemner, C., Dietz, C., Swinkels, S. H. N., Buitelaar, J. K., &

    Van Engeland, H. (2009). Inter-rater reliability and stability of diagnoses

    of autism spectrum disorder in children identified through screening at a

    very young age. *European Child and Adolescent Psychiatry, 18*, 663–

    674. doi:10.1007/s00787-009-0025-8

Vaughan, C. A. (2011). Test Review: E. Schopler, M. E. Van Bourgondien, G.

    J. Wellman, & S. R. Love Childhood Autism Rating Scale (2nd ed.). Los

    Angeles, CA: Western Psychological Services, 2010. *Journal of*

    Psychoeducational Assessment, 29, 489-493.

    doi:10.1177/0734282911400873

Ventola, P., Kleinman, B., Pandey, J., Barton, M., Allen, S., Green, J., . . . Fein,

    D. (2006). Agreement among four diagnostic instruments for autism

    spectrum disorders in toddlers. *Journal of Autism and Developmental*

    Disorders, 36, 839–847. doi:10.1007/s10803-006-0128-8

Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology, 56*, 315–336. doi:10.1146/annurev.psych.56.091103.070159

Volkmar, F. R., Klin, A., Siegel, B., Szatmari, P., Lord, C., Campbell, M., . . . & Towbin, K. (1994). Field trial for autistic disorder in DSM-IV. *The American Journal of Psychiatry, 151*, 1361-1367.

Volkmar, F., Paul, R., Rogers, S. J., & Pelphrey, K. A. (2014). *Handbook of Autism and Pervasive Developmental Disorders, Assessment, Interventions, and Policy. Volume 2 of Handbook of Autism and Pervasive Developmental Disorders* (4th ed.). Hoboken, NJ: John Wiley & Sons

Volkmar, F., Siegel, M., Woodbury-Smith, M., King, B., McCracken, J., State, M., American Academy of Child and Adolescent Psychiatry (AACAP), & Committee on Quality Issues (CQI). (2014). Practice parameter for the assessment and treatment of children and adolescents with Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*, 237-257. doi:10.1016/j.jaac.2013.10.013

Wallis, K. E., & Pinto-Martin, J. (2008). The challenge of screening for autism spectrum disorder in a culturally diverse society. *Acta Paediatrica, 97*, 539–540. doi:10.1111/j.1651-2227.2008.00720.x

Walsh, P., Elsabbagh, M., Bolton, P., & Singh, I. (2011). In search of biomarkers for autism: scientific, social and ethical challenges. *Nature Reviews Neuroscience*, *12*, 603-612. doi:10.1038/nrn3113

Wang, H., Chen, C., & Fushing, H. (2012). Extracting multiscale pattern information of fMRI based functional brain connectivity with application on classification of autism spectrum disorders. *PLoS One, 7*, e45502. doi:10.1371/journal.pone.0045502

Watkins, M. W. (2000). *Monte Carlo PCA for Parallel Analysis* [computer software]. State College: Ed & Psych Associates.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children (4th ed.).* San Antonio, Texas: Psychological Corp.

West, S. S., Flinch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56-75). Thousand Oaks, CA: Sage.

Wiggins, L. D., Bakeman, R., Adamson, L. B., & Robins, D. L. (2007). The utility of the Social Communication Questionnaire in screening for Autism in children referred for early intervention. *Focus on Autism and Other Developmental Disabilities, 22*, 33-38. doi:10.1177/10883576070220010401

Wilkins, J. & Matson, J. L. (2009). A comparison of social skills profiles in intellectually disabled adults with and without ASD. *Behavior Modification, 33*, 143-155. doi:10.1177/0145445508321880

Wing, L. (1996). *The Autistic Spectrum: A Guide for Parents and Professionals*. Constable, London.

Witwer, A. N. & Lecavalier, L. (2007). Autism screening tools: An evaluation of the Social Communication Questionnaire and the Developmental Behaviour Checklist-Autism Screening Algorithm. *Journal of Intellectual & Developmental Disability Research, 32*, 179–187. doi:10.1080/13668250701604776

World Health Organization (1993). *The ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research*. WHO, Geneva.

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*, 965–970. doi:10.1016/j.rasd.2011.12.012

Young, R. L. (2007). *Autism Detection in Early Childhood: ADEC*. Australian Council of Educational Research, Camberwell, VIC, Australia.

Young, R. L., Brewer, N., & Pattison, C. (2003). Parental identification of early behavioural abnormalities in children with autistic disorder. *Autism, 7*, 125–143. doi:10.1177/1362361303007002002

Young, R. L., Partington, C., & Goren, T. (2009). *The structured program for early childhood therapists working with Autism (SPECTRA)*. Australian Council of Educational Research, Camberwell, VIC, Australia.

Young, R. L., & Rodi, M. (2014). Redefining Autism Spectrum Disorder using DSM-5: The implications of the proposed DSM-5 criteria for Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 44*, 758-765. doi:10.1007/s10803-013-1927-3

Zwick, W. R. & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.

APPENDIX A

Details of ADEC Items and Scoring Protocols

ADEC Item	Operationalisation	Scoring Protocol
1. Response to Name	**Example**: • Child turns head and looks at tester's face and makes eye contact **Non-examples**: • Child does not look up from activity • Child looks around but not at tester's face	**0**: child turns head towards tester immediately following name call on first or second trial **1**: child turns head towards tester immediately following name call on third, fourth or fifth trial; or behaviour is seen to occur spontaneously at other times during the testing session **2**: child does not respond to name on any of the 5 trials and this behaviour is not demonstrated spontaneously during the testing session
2. Imitation (drum hands on box)	**Example**: • Child drums on box with both hands **Non-examples**: • Child drums on box with just one hand • Child does not respond • Child looks away	**0**: child drums on box with both hands on at least one trial **1**: child makes clear attempt to imitate the gesture but is impeded by lack of motor co-ordination or some spontaneous imitation occurs during testing but not on command (score '1' if child imitates any actions during the testing session **2**: child makes no attempt to imitate gesture on any of the 3 trials
3. Stereotypical Behaviour (upset when line of blocks disturbed)	**(a) Child becomes distressed when the blocks are disturbed** **Example**: • Child cries or screams **Non-examples**: • Child disturbs the line-up of blocks on their own initiative • Child does not respond to the	**0**: child unconcerned by disturbance of line-up of blocks or disturbs them him or herself **1**: child becomes upset at disturbance of line-up of blocks or demonstrates some linear alignment of other objects such as cars during the adaptation or testing sessions **2**: child becomes upset *and* attempts to realign the blocks

| | disturbance of the line of blocks

(b) Child makes attempts to realign the blocks

Example:
- Child attempts to place one or more of the moved blocks back into a linear arrangement
- Child builds blocks in some order (colour) and is upset when it is disturbed

Non-examples:
- Child disturbs the line of blocks on their own
- Child does not respond to the disturbance of the line of blocks
- Child starts to align the blocks in a non-linear fashion (e.g., building a tower in a non-structured format)
- Spontaneous alignment of objects other than blocks or some rigidity in positioning of objects should be scored as '1' here. That is, if the child makes any attempt to line up or stack objects other than the blocks, or is disturbed by the tester's repositioning of objects placed by him/her this should be | |

	scored as '1' here. In addition, any stereotypical behaviours (including body movements) should score '1'.	
4. Gaze Switching	**Example**: • Child points at toy and also looks at adult's face (either caregiver's or tester's) • Child turns head and eyes to look at toy then turns head and eyes to look at the tester's face (and back at the toy again) • Child looks at adult's face **Non-examples**: • Child does not look up • Child becomes upset (cries) without looking at adult's face • Child is completely uninterested	**0**: child turns head and eyes to look at toy then turns head and eyes to look at adult, then turns back to look at toy; (may combine eye gaze with pointing/vocalisations/reaching) **1**: child may look at either toy or adult but with no gaze switching between the toy and adult (i.e., he/she makes no attempt to look or engage adult) **2**: child makes no attempt to look at or engage adult; child may be just fixated on toy and indifferent to surroundings or may be indifferent to toy
5. Eye Contact (in a game of peek-a-boo)	**Example**: • Child engages in game and shows good eye contact • Child displays signs of interest in the game (e.g., child smiles; child laughs) • Child becomes excited and looks in the tester in the eye	**0**: on each of the 5 trials, child engages in game and looks into tester's eyes **1**: on at least one, but not all 5 trials, child looks into tester S eyes **2**: child does not look into tester's eyes on any of the 5 trials or deliberately tries to avoid eye contact

	Non-examples: • Child displays signs of no interest in the game • Child looks away; child does not engage in eye contact	
6. Functional Play (toy telephone)	**Example**: • Child pushes car along • Child picks up receiver and holds it to their ear • Child picks up receiver and holds it to their ear and vocalises • Child dials telephone **Non-examples**: • Child plays with only one feature of the toy (e.g., spinning the wheels) • Child engages in sensorimotor play (e.g., banging, waving, sucking, throwing, sniffing)	**0**: child engages in *more than one* of the Example behaviours **1**: child engages in *only one* of the Example behaviours throughout the time period/or functional play was observed throughout the session using other toys **2**: child does not engage in any of the Example behaviours
7. Pretend Play (pretend phone)	**Example**: • Child holds the piece of foam to ear, as if it is a telephone receiver **Non-examples**: • Child takes the piece of foam from tester but does not hold it against ear (e.g., holds it, bangs it against table, throws it, eats it)	**0**: child holds foam to ear, as if it is a telephone receiver and vocalises **1**: child takes phone, maybe he/she vocalises but does not clearly demonstrate an understanding that it is a pretend phone **2**: child displays any of the Non-example behaviours

	• Child does not take the piece of foam from tester • Child looks away from the tester	
8. Reciprocity of Smile	**Example**: • Child smiles **Non-examples**: • Child looks at tester but does not smile • Child looks away from tester's face	**0**: child smiles immediately after one of first 2 trials, and there is a clear change in expression from a non-smiling expression to smile **1**: delayed smile or smile occurs spontaneously during the testing session **2**: child does not smile; avoids social contact with tester throughout the testing session
9. Response to Everyday Sounds	**Example**: • Child turns head towards CD player • Child points to CD player • Child looks at caregiver or tester **Non-examples**: • Child covers ears with hands • Child attempts to remove self • Child cries or screams	**0**: child engages in any of the Example behaviours and there are no Non-example behaviours demonstrated **1**: child ignores sound, continues with his or her activity **2**: child engages in any one of the Non-example behaviours
10. Gaze Monitoring (following point/pointing)	**Example**: • Child turns head to look in the direction tester is pointing • Child points to something in the room **Non-examples**: • Child looks at tester's face, hand or arm but does not follow point or point themselves	**0**: child turns head and looks in the direction tester is pointing or child points to something spontaneously to engage tester **1**: child does not look at object of interest, instead focuses on the tester (i.e., the tester's face, or pointing hand or arm) **2**: child does not look up at the tester; child looks away or no pointing is observed

	• Child does not respond	
11. Response to a Verbal Command	**Example**: • Child responses appropriately to a verbal command (e.g., 'clap hands', 'blow kiss')	**0**: child demonstrates the behaviour to the standard normally achieved according to the caregiver **1**: child responds to the command (looks up at caregiver, approaches caregiver) but does not demonstrate the behaviour he/she was asked to do **2**: child does not respond or looks away from caregiver; or caregiver states that child is unable to respond to a verbal command
12. Demonstrates Use of Words	**Example**: • For child 12-18 months of age: child demonstrates use of at least one word, clearly pronounced, and is not a made-up word • Child demonstrates at least 6 words (18 months to 2 years) • Child demonstrates more than 12 words (more than 2 years)	**0**: child clearly pronounces one word or more (12-18 months of age); child demonstrates at least 6 words (18 months to 2 years); child demonstrates more than 12 words (more than 2 years of age) **1**: child makes an attempt but the word is not pronounced clearly; or child just babbles; or fewer words are spoken than is desirable for the child given their age **2**: child does not use any words
13. Anticipatory Posture (for being picked up)	**Example**: • Child leans forward towards caregiver • Child raises one or both elbows/arms to make armpits available for caregiver to grasp • Child displays anticipation of being picked up but appears to reject it –	**0**: child demonstrates one of more of the Example behaviours making it clear they realise the intent of the caregiver **1**: child displays one of more of the Example behaviours after much prompting (either verbal or physical) **2**: child does not display any of the Example behaviours

	this might be by squeezing his/her arms against their own body as if to prevent the caregiver from gaining access to their armpits **Non-examples**: - Child looks at caregiver but does not raise arms - Child looks away from caregiver or child continues activity without responding	
14. Nestling into Caregiver	**Example**: - Child nestles into caregiver by resting body on caregiver's body and leaning head on caregiver's shoulder (may put arms around caregiver) **Non-examples**: - Child assumes rigid posture - Child assumes limp posture (needs to be held up by caregiver) - Child struggles - Child arches back - Child pushes caregiver away	**0**: child displays the Example behaviour (nestling into caregiver) **1**: child displays some indications of discomfort but none of the Non-example behaviour; or child will only nestle at their initiative not when responding to parent/caregiver; or behaviour is seen to occur spontaneously at other times **2**: child displays one or more of the Non-example behaviours

15. Use of Gestures (wave goodbye)	**Example**: • Child waves at tester at least once while waving (arm or hand is extended towards tester and waved side to side and/or up and down repeatedly, or hand is opened and closed, palm facing towards the tester **Non-examples**: • Child extends arm towards tester but does not move it up and down in a waving action • Child does not respond	**0**: child displays the Example behaviour **1**: child makes clear attempt to wave (e.g., child extends arm towards tester but does not move it up and down in a waving action). Any other demonstration of these types of gestures can be scored here **2**: child does not respond or looks away
16. Ability to Switch from Task to Task	**Example**: • Child shifts from one task to another with little resistance **Non-examples**: • Child becomes fussy if activity is changed • Child becomes fixated with one particular task • Child does not engage sufficiently in any activity that enables a shift to be observed	**0**: child readily changes from one activity to another **1**: child may become fixated on one task but generally is happy to change tasks **2**: child does not respond to tester's requests to change tasks or will not engage in tasks as required preferring to do their own thing

APPENDIX B

Autism Detection in Early Childhood (ADEC) Score Sheet

ADEC Score Sheet

CHILD'S NAME_____   ID No_____

Child's date of birth: _____

Tester _____ Date of testing :_____

Observations of Behaviour:

Item		Appropriate		Inappropriate
1	Response to name	0	1*	2
2	Imitation (drum on box)	0	1	2
3	Upset when line of blocks is disturbed	0	1	2
4	Gaze-switching (Tigger or other toy)	0	1	2
5	Eye-contact in game of Peek-a-boo (engagement)	0	1	2
6	Functional play (toy telephone/car)	0	1	2
7	Pretend play (pretend phone)	0	1	2
8	Reciprocity of a smile	0	1	2
9	Response to everyday sounds	0	1	2
10	Gaze monitoring – follow point	0	1	2
11	Response to verbal command	0	1	2
12	Demonstrates use of words	0	1	2**
13	Anticipatory posture for being picked up	0	1	2
14	Nestling into caregiver	0	1	2
15	Use of Gestures (wave goodbye)	0	1	2
16	Ability to switch from task to task	0	1	2
	Total Score	_____		

*Score "1" if child spontaneously demonstrates behaviour but not when required. on any of the items.

** Score 2 if child is: older than 2 years of age and has less than 12 words
between 18 months and 2 years and has less than 6 words
between 12 months and 18 months and has less than one word.
less than 12 months and is not babbling

Score	Risk of Autism	Possible Follow-up
0-10	Low-Risk	No immediate action required
11-13	Moderate-Risk	Child should be reviewed.
14-19	High-risk	Further testing required
>19	Very High-risk	A formal autism assessment is strongly recommended.