Examining the clinical utility of the modified Alarm Distress Baby Scale (m-ADBB) for the detection of early signs of Autism Spectrum Disorder

Marion Bettina Christl

The University of Sydney | Faculty of Science| School of Psychology

SID: 200341556

Declaration

This thesis is submitted to the University of Sydney in fulfilment of the requirements for the Master of Science.

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Marion Bettina Christl

Signature: *B. Christl*

Date: 19/12/2016

Acknowledgements

Table of Contents

Index of Table and Figures

Tables

Figures

Abstract

Background / Aim

Although behavioural signs of Autism Spectrum Disorder (ASD) can be observed within the first two years of life, early detection and diagnosis of ASD remains challenging. Existing routine screening instruments have significant limitations, either being too time-consuming to use or not being reliably validated for routine screening. Therefore, no reliable and easy to use routine screening instrument for ASD in infants and toddlers is currently available to clinicians.

Some of the early ASD related deficits in social communication and interaction behaviours are also typical of sustained social withdrawal in infants and toddlers. Only one study has investigated this link to date, and its findings suggest that sustained social withdrawal in infants and toddlers may be indicative of ASD (Wendland, Gautier, Wolff, Brisson, & Adrien, 2010).

The current study aims to test whether a brief observational screening instrument for social withdrawal in infants, the modified Alarm Distress Baby Scale (m-ADBB) may be clinically useful for the detection of ASD in the first two years of life. It is hypothesised, that children with ASD will score higher on the m-ADBB than typically developing (TD) children, indicating more symptoms of social withdrawal.

Method

Home-video recordings of children with ASD and children with typical development from approximately age 12 months and 24 months were analysed using the m-ADBB.

Results

Eleven children with a diagnosis of ASD and eleven children with typical development were recruited to the study; videos for 20 children were available at each age. Children with a diagnosis of ASD scored statistically significantly higher on the m-ADBB than children with typical development at 12 month (Z=-2.54; p=0.023; r=-0.57) and at 24 month (Z=-2.40; p=0.023; r=-0.54). Five of ten children with ASD met the m-ADBB criterion for social withdrawal in their 12 month videos, and four out of ten in their 24 month videos. Using a lower cut-off score increased detection rates (7 at 12 month; 8 at 24 month). False positive results were low at both ages and when both cut-offs were applied (range 1 to 3 out of 10).

Conclusion

The findings suggest that observing only five social withdrawal behaviours as operationalized by the m-ADBB is useful in flagging possible presence of ASD in children during their first two years of life. Further research is required to establish the scale's sensitivity and specificity for ASD detection.

Glossary

AD	Autistic Disorder
ADI-R	Autism Diagnostic Interview Revised
ADOS	Autism Diagnostic Observation Schedule
ASD	Autism Spectrum Disorder
BITSEA	Brief Infant-Toddler Social and Emotional Assessment
CRS	Clinical Rating Scale
DD	Developmental delay
DSM	Diagnostic and Statistical Manual of Mental Disorders
ESAT	Early Screening for Autism Traits
FYI	First Year Inventory
ID	Intellectual disability
ITC	Infant-Toddler Checklist
M	Mean
m-ADBB	Modified Alarm Distress Baby Scale
M-CHAT	Modified Checklist for Autism in Toddlers
NPV	Negative Predictive Value
PDD-NOS	Pervasive Developmental Disorder - Not Otherwise Specified
PEDS	Parent Evaluation of Developmental Status
PPV	Positive Predictive Value
SACS	Social Attention and Communication Study
SD	Standard deviation
SRS	Social Responsiveness Scale
STAT	Screening Tool for Autism in Two-Year-Olds
TD	Typical Development/ typically developing
TSC	Time-sampled coding

Introduction

Autism Spectrum Disorder (ASD) is a pervasive developmental disorder characterised by impaired social interaction and communication abilities, leading to difficulties in personal relationships, education and employment, as well as general independence across the life span (Seltzer, Shattuck, Abbeduto, & Greenberg, 2004). Research shows that intervention in early childhood can remediate the severity of impairments (Dawson, 2008; Dawson et al., 2010; Matson, 2007) and improve functioning in later life (Howlin, 1997). To reap the full benefit of early intervention, early detection of ASD is essential, but to this day this remains challenging; no physiological or genetic test is available for ASD at this stage, and detection relies solely on observation of early behavioural signs. There is currently no reliable or easy to implement ASD screening test available for infants and toddlers, and this contributes to considerable delay in diagnosis and intervention for many children.

To illustrate the delay in diagnosis, in Australia, on average parents notice ASD symptoms around age nine month but professional investigations on average do not start until 12 months later (Young, Brewer, & Pattison, 2003). The complexity of a diagnostic ASD assessment means that it can take another year or more until a diagnosis is confirmed (Wiggins, Baio, & Rice, 2006). The average age of diagnosis of ASD in Australia is 4 years, with many children being diagnosed only when entering the education system (Bent, Dissanayake, & Barbaro, 2015). Consequently, many families endure prolonged uncertainty and anxiety about their child's development despite early manifestation of symptoms, and many children are missing out on the benefits of early intervention. Thus the development of a reliable ASD screening instrument for use in routine clinical practice is a meaningful and important endeavour.

The present study aims to examine whether an existing validated observational scale for the assessment of social withdrawal in infants and toddlers may be clinically useful for the early detection of ASD. To address these aims, this thesis provides a review of the relevant literature (Chapter 1), it includes a systematic review of studies employing home-video analysis to examine the early signs of ASD (Chapter 2), it further describes the methodology used for the empirical study (Chapter 3), presents the findings of the home-video analysis (Chapter 4), and concludes with an overall discussion of the research findings, the theoretical and clinical implications, and the strengths and limitations of the current study (Chapter 5).

CHAPTER 1:   Literature Review

1.1   What is Autism Spectrum Disorder?

1.1.1   Diagnostic Criteria of Autism Spectrum Disorder

Autism Spectrum Disorder is a pervasive developmental disorder present from birth. The diagnostic criteria for autism spectrum disorder and its related conditions have changed and been refined multiple times since autism was first described at the beginning of the 19th century by Leo Kanner (Happé & Ronald, 2008). The diagnostic criteria underwent their most recent change in the DSM-5, the latest revision of the Diagnostic and Statistical Manual of Mental Disorders (APA, 2013). The DSM-5 has introduced the diagnostic label Autism Spectrum Disorder, subsuming the previously separate four diagnostic categories of Autistic Disorder, Asperger's Syndrome, Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) and Childhood Disintegrative Disorder. The grouping of these diagnoses into one diagnostic category reflects the already longstanding use of the term Autism Spectrum Disorder in research and clinical practice (Carroll, 2015).

According to the DSM-5, ASD is characterised by persistent deficits in social communication and social interaction across multiple contexts as well as restricted, repetitive patterns of behaviour, interests, or activities (APA, 2013). These symptoms have to be present in "the early developmental period"; however the DSM-5 acknowledges that symptoms may not manifest clearly "until social demands exceed limited capacities"(APA, 2013).

Social communication and interaction deficits are defined as deficits in in social-emotional reciprocity, nonverbal communication and deficits in developing, maintaining, and understanding relationships (APA, 2013). The DSM-5 recognises impaired joined attention as one of the very early signs of these deficits, which manifests in lack or reduced use of gestures and gaze in drawing someone else's attention to an object or event, or following someone else's gestures and gaze to join them in their focus of attention.

Restricted, repetitive patterns of behaviour, interests, or activities are defined as stereotyped or repetitive motor movements, use of objects, or speech; this also includes insistence on sameness, rigid adherence to routines, or ritualistic verbal or nonverbal behaviour; furthermore it can manifest as highly restricted, fixated interests that are abnormal in intensity or focus and sensory processing issues, such as hyper- or hypo-reactivity to sensory input or seeking out specific sensory stimulation (APA, 2013).

1.1.2 Prevalence of Autism Spectrum Disorder

Prevalence rates for ASD have changed over the years, due to the changes in diagnostic criteria. Most recent available estimates are based on the DMS-IV criteria. In Australia, the combined prevalence for Autistic Disorder, Asperger's Syndrome and PDD-NOS as defined by DSM-IV–TR (APA, 2000), is estimated to be 1 in 160 (Barbaro & Dissanayake, 2010); this is slightly lower than prevalence rates of about 1% found in international studies (Fombonne, 2003). The prevalence rate of Autistic Disorder alone, the most severe form of ASD, is about 4.3/10,000 (Williams et al., 2005). AD, PDD-NOS and Asperger's Syndrome together have been reported to be 4 or 5 times more prevalent among males than females; however, the gender gap is most pronounced in high functioning individuals (IQ ≥ 70) and least in individuals with co-occurring intellectual disability (Lai, Lombardo, Auyeung, Chakrabarti, & Baron-Cohen, 2015). It has been suggested that the higher male prevalence is due to the under-recognition of females in the higher functioning category and gender biased diagnostic criteria and assessment instruments (Lai et al., 2015); this under-recognition subsequently leads to females being later diagnosed than males (Begeer et al., 2013; Giarelli et al., 2010). Epidemiological studies accounting for these factors still find 2- 5 times higher prevalence rates for males than females for high-functioning ASD (Lai et al., 2015). Prevalence rates have been reported to have risen during the time the DSM-IV diagnostic criteria applied, mainly within the high-functioning group (Lai et al., 2015) potentially due to increased awareness and improved recognition of the condition. However, it is expected that the revised DSM criteria will revert this trend as the new criteria are seen as being stricter than the previous ones (Matson, Kozlowski, Hattier, Horovitz, & Sipes, 2012; Mattila et al., 2011; McPartland, Reichow, & Volkmar, 2012).

1.1.3 Onset Patterns of Autism Spectrum Disorder

The onset of ASD has long been thought to be either progressive with early signs from birth or regressive with loss of skills, typically occurring around 18 months; this is also referred to as late onset.

The distinction of progressive and regressive onset has first been made by Japanese researchers (see Kobayashi and Murata (1998)) and since been used widely in subsequent research studies. The concept of regressive or late onset has important implications for the early detection of ASD; screening before a certain age (i.e. 18 months) may not detect children with regressive or late onset. However, the validity of the concept of late or regressive onset has been questioned recently. Firstly, rates of regressive onset have been varying greatly from 15% up to 50% in ASD samples due to varying definitions of what constitutes regression (Ozonoff, Heung, Byrd, Hansen, & Hertz-Picciotto, 2008a). If regression is defined to include loss of language skills around 15% children with ASD are considered to have this onset pattern (Ozonoff et al., 2008a).

However, in their review of the evidence about ASD onset patterns, Ozonoff and colleagues (2008) argue that the dichotomised onset classification in early onset versus regressive or late onset may not be exhaustive for how ASD manifests across the spectrum; instead they suggested a third onset pattern which is characterised by an intact early social development, with possibly non-specific abnormalities being present, followed by a failure to progress in development (developmental plateau). This third onset pattern has been confirmed by two observational home-video studies (Ozonoff et al., 2011; Wendland et al., 2010). Furthermore, when home-video observations were compared with parent report about onset of ASD symptoms, it was found that atypical development was indeed present in many children with ASD during their first year without their parents recognising this. Parents typically noticed the more salient loss of language or other skills in the second year, thus giving the impression of a late onset when in fact other symptoms were already present during the first year. This finding suggests that parent report may not be a reliable source of evidence to confirm the regressive or late onset pattern (Ozonoff et al., 2011). It further questions the validity of rates of regressive onset reported in previous studies which relied on retrospective parent report about onset and type of symptoms in their children.

The debate about onset patterns has progressed further, and the possibility of four onset patterns has been suggested (Ozonoff et al., 2008a). These onset patterns include: 1) an early or progressive onset of symptoms from birth; 2) no early signs and abrupt loss of skills or regression later; 3) early signs and additional later loss of skills; or 4) an early intact or normal development with a later plateau in development. Alternatively it has been suggested that the evolvement of ASD symptoms may be even more heterogeneous and thus better be described by an onset continuum, with symptoms from birth at one end and an abrupt loss of skills without any prior symptoms at the other end (Ozonoff et al., 2008a).

In conclusion, more recent research suggests that some reports of regressive onset may be an artefact of parent recall, and that it may occur much less frequently than previously thought. Onset patterns may also be potentially more varied than the dichotomisation into progressive or early onset and regression or late onset. For early detection of ASD this suggests that most children are expected to show some observable behavioural differences in the first year of life, but that a small proportion of children may not show any signs until after their second birthday, and thus these children may be missed even if early screening was available.

1.1.4 Autism Spectrum Disorder and Intellectual Disability

ASD is highly comorbid with intellectual disability (ID). Across the entire ASD spectrum it is estimated that 50–70% of individuals with ASD also meet criteria for ID (Matson & Shoemaker, 2009). The high rate of comorbid ID adds to the difficulties of diagnosing ASD early, because some of the early signs of impairments in social communication are common to both ID and ASD, and thus they could be falsely attributed as symptoms exclusively of ID

rather than ASD. On the other hand, parental concerns about early signs of ASD seem to be triggered earlier for children with comorbid ID (see Chawarska et al. (2007)) which in turn may make early detection of ASD more likely for children with comorbid ID as investigations may commence earlier. Because of the high comorbidity rate of ID, ASD without the presence of ID is often referred to as high-functioning autism in the literature.

1.2 Early Signs of Autism Spectrum Disorder

1.2.1 Research methodologies
Knowledge about early behavioural signs of ASD is the basis on which early screening tools and procedures are developed. During the past decades extensive research efforts have been made to ascertain the behavioural signs in the first two years of life that may be indicative of ASD. Research to date has employed different methodologies to investigate early ASD signs. The different methodologies are described below.

High-risk sibling studies follow the development of younger siblings of children with a diagnosis of ASD; these are regarded to be at high-risk for ASD due to genetic loading (Grønborg, Schendel, & Parner, 2013). 'High-risk' siblings and a comparison group of 'low-risk' children, i.e. children without older siblings with ASD, are typically followed until a diagnostic assessment at two or three years of age (Mitchell, Cardy, & Zwaigenbaum, 2011; Rogers, 2009). The observations are typically conducted in standardised lab settings, which may be limiting the ecological validity of their findings. More recently high-risk sibling studies have started to investigate neurological correlates of ASD through use of sophisticated technology, such as eye-tracking, EEG and MRIs (see Costanzo et al. (2015)). These studies are less suited to identify early signs of ASD that can be easily observed in clinical practice.

Studies using retrospective parent report (Barbaro & Dissanayake, 2009; Mitchell et al., 2011; Zwaigenbaum et al., 2005) assess the presence of early signs through questionnaires or interviews with parents of children who already have been diagnosed with ASD. This method has the advantage that parents may draw from a wide range of experiences with their child rather than relying on a standardised observational situation as it is the case in high-risk sibling studies. However, parent report may be inaccurate due to recall bias or limited parental knowledge of ASD symptoms or typical development (Zwaigenbaum et al., 2005); the inaccuracy of parent report regarding onset of symptoms has already been discussed earlier in this chapter.

Retrospective analysis of home-video recordings (Palomo, Belinchón, & Ozonoff, 2006; Saint-Georges et al., 2010) is another approach for investigating early behavioural symptoms of ASD. This approach has been used by numerous research groups in the USA, Italy, France, Germany and Australia (Costanzo et al., 2015; Mitchell et al., 2011; Palomo et al.,

2006; Saint-Georges et al., 2010). The advantage of home video studies is that they do not require a large sample size because the diagnostic status of the participants is already known. Importantly, they allow for objective observations without having to rely on parents' retrospective report. Home-videos analysed for this purpose are typically recorded by parents during the first two years of life of their child. Behaviours observed by researchers in these videos are then compared with those observed in videos of typically developing children and/or children with other developmental delay at matched ages. While this methodology is more feasible than prospective study designs and more objective than parent report studies, it is not without its limitations. Home-video studies lack standardisation of the context in which a child is observed and the recordings vary in length and quality. Home-video recordings also do not represent random samples of a child's behaviour but rather particular occasions such as birthdays, holidays or developmental milestones, and thus are subject to selection bias by the parents; e.g. parents may record and share only those videos with researchers that show their child at their best behaviour and chose to turn off the camera once their child shows undesirable behaviours; or vice versa, parents may select videos that show behaviours that are of most concern to them in order to have clinicians and researchers assess these. All these factors however, can be assessed and their impact on study outcomes controlled for. Despite these limitations, several reviews of home-video studies have shown that this methodology has scientific merit in examining early behavioural signs of ASD (Barbaro & Dissanayake, 2009; Costanzo et al., 2015; Palomo et al., 2006; Saint-Georges et al., 2010). And this research methodology has been chosen for the current study.

Several systematic and non-systematic reviews of high-risk sibling, parent report and retrospective home-video studies have been published in the recent years (Barbaro & Dissanayake, 2009; Costanzo et al., 2015; Mitchell et al., 2011; Palomo et al., 2006; Rogers, 2009; Saint-Georges et al., 2010); an overview of the findings of these reviews are given in this following section. Because the social interaction and communication behaviours are of particular relevance to the current study, the overview will focus on these and will make broader comments on the evidence for other behavioural signs. A systematic review of the evidence from home-video studies will be offered in Chapter 2.

Generally, synthesising the evidence for early signs of ASD is challenging for a variety of reasons. The different study methodologies have been described already, and these may contribute to inconsistencies in study findings and their validity. Studies also vary in the types of behaviours they investigated. Furthermore, the diagnostic categories included and differentiated in these studies vary, i.e. some studies include children from the whole autism spectrum, while some include only those with a diagnosis of autistic disorder and / or PDD-NOS; Asperger's Syndrome has been excluded from most studies. Studies also differ in whether they compare children with ASD to typically developing (TD) children only or additionally to children with other developmental delay (DD). The latter comparison

provides insight into behavioural differences that are specific to ASD and those that are indicative of any developmental difficulties, and thus this comparison provides important information for the development of ASD specific screening tools. Furthermore, different studies have divided the first two years of life into different not always comparable age periods. And lastly, different studies focus on different behaviours and measure these behaviours in different ways (presence/absence; frequency; quality). For the purpose of this review, evidence has been considered to be consistent if it is supported across different study methodologies, and if the number of studies supporting the evidence is greater than the number of studies not supporting it.

1.2.2 Early social interaction and communication deficits in infants and toddlers with ASD

The evidence for early deficits in social interaction and communication behaviours will be presented for the first six months, then for the six to 12 months period and lastly for the second year, as this follows the most common grouping of ages in the existing studies.

Age zero to six months

A review of high risk sibling studies (Jones, Gliga, Bedford, Charman, & Johnson, 2014) found that children later diagnosed with ASD were no different than TD children during the first six months of their life in regard to social smiling and social vocalisations; infants with ASD have further been found to look at faces with typical patterns of face scanning and eye contact; they also showed shared affect and interest in others and were overall as socially responsive as TD children. There is, however, evidence from retrospective home-video studies, that children with ASD show less social attention, less anticipation of others' aim, impaired pointing comprehension and attuning behaviours as well as reduced exploratory activity, when compared to typically developing children (Saint-Georges et al., 2010). None of the home-video studies investigating the first six months have used a comparison group of children with other developmental delay; thus it is unclear if the behavioural differences found at this age are ASD specific or whether they are signs of general developmental abnormality.

The inconsistency between findings for the first 6 months from home-video studies and high-risk sibling studies indicates that identification of symptoms during the first six months is challenging. This may also be due to different onset patterns of ASD in the examined samples, meaning that some signs may only be observable in the first 6 months in children with early onset from birth. No parent report studies have investigated social interaction and communication behaviours in the first 6 months. Given the evidence for this very early age period, developing a reliable screening tool for ASD before the age of 6 months may not be possible.

Age six to twelve months
Use of Gestures and joint attention behaviours. Joint attention is the ability to share attention to the same object or event with another person, either by initiating or responding to joint attention behaviours, such as showing objects, pointing, head turning, gaze alternation between a person and an object and vocalisations; often these behaviours are used in combination. Some studies have examined the use of gestures alone as form of non-verbal communication, while other studies have examined them as part of other joint attention behaviours.

Evidence from high-risk sibling studies suggests that children with ASD use less communicative gestures (i.e. pointing, showing, waving, nodding or shaking head) than TD children at age 6 to 12 months; children with ASD were found to less frequently point or show objects to others, and they had a smaller repertoire of gestures overall. They were also less likely to us pointing gestures to initiate joint attention (see Jones et al. (2014)). High-risk siblings who later were diagnosed with ASD have also been found to be less likely to combine their gaze with gestures and vocalisations to engage others, however, gaze alternation alone was not different at this age in children with ASD when compared to DD and TD children (see Jones et al. 2014). Parent report studies support these findings of reduced use of gestures and reduced use of combination of looking and pointing at 12 months (see Jones et al. (2014)). Home-video studies also consistently found that children with ASD used fewer communicative gestures; when compared to TD children by age 12 months (Saint-Georges et al., 2010). A recent home-video study, not included in previous reviews, found that joint attention behaviours, i.e. combining gestures and gaze, was also reduced in ASD infants when compared to TD and DD infants (Watson, Crais, Baranek, Dykstra, & Wilson, 2013). Overall findings across study methodologies consistently found deficits in joint attention behaviours and use of gestures. But more research is needed to confirm whether this is indeed an ASD specific marker in the first year or rather a sign of general developmental abnormality.

Eye contact. The evidence on deficits in eye contact of children with ASD in the first year is limited to findings of only four studies, two home-video and two high risk sibling studies; of these, one home-video study found a lower frequency and quality of eye contact in children with ASD compared to TD children (Saint-Georges et al., 2010), and one high-risks sibling study found atypical eye-contact at 12 months (Jones et al., 2014); the other two studies found no differences between children with ASD and TD children. Comparing birthday videos to non-birthday videos, another home-video study found that at 12 months children with ASD displayed significantly more eye contact and less vocalisation in birthday videos than in non-birthday videos (Thorsen, Goldberg, Osann, & Spence, 2008). This highlights the importance of accounting for the context in which behaviours are observed. Overall impaired or reduced eye contact appears to not be a reliable indicator of ASD in the first year.

Affect behaviours. Home-video studies consistently found social smiling as not to be reduced in children with ASD in the first year (Saint-Georges et al., 2010). However, high-risk sibling studies found that children with ASD showed less social smiling than TD children (Jones et al., 2014). Home-video studies furthermore found that children with ASD showed fewer facial expressions or that these were less appropriate in infants with ASD when compared to TD infants (Palomo et al., 2006; Saint-Georges et al., 2010). Overall home-video studies found reduced frequency and quality of affective expression to be an ASD specific early sign (Saint-Georges et al., 2010). Evidence from one high-risk sibling study found that children with ASD show reduced affective responsiveness t (see Jones et al, 2014).

Response to name. Both high-risk sibling studies and home-video studies found that children with ASD were less responsive to their name being called than TD children and children with other DD (Barbaro & Dissanayake, 2009; Mitchell et al., 2011), and thus, this can be seen as an ASD specific early sign (Saint-Georges et al., 2010).

Social attention and interest. Home-videos studies also found children with ASD looked less at other people and more at objects (Mitchell et al., 2011; Palomo et al., 2006; Saint-Georges et al., 2010); one home-video study investigated the development of social attention behaviours (smiling, looking, orienting and vocalising to people) and found that by 12 months children with ASD preferred to pay attention to objects over people (Maestro et al., 2005b).

Vocalisation. High-risk sibling studies found mixed evidence for early delays in expressive and receptive language for children with ASD in the first year of life (Jones et al., 2014). Home-video studies found that children with ASD did not overall have fewer vocalisations when compared to TD children but that they had fewer vocalisations directed at others (Palomo et al., 2006; Saint-Georges et al., 2010).

Table 1.3.2 provides an overview of the signs consistently found to distinguish ASD from TD and DD in the first year.

Age 12- 24 months:
Use of Gestures and joint attention behaviours. High-risk sibling studies, home-video studies and parent report studies all consistently reported that children with ASD use fewer gestures than TD children in the second year of life and initiate less joint attention with others (Jones et al., 2014; Saint-Georges et al., 2010). High-risk sibling studies also found that children with ASD showed reduced response to joint attention actions of others (Jones et al., 2014; Rogers, 2009), such as following another person's gaze, head turn, verbal cues or pointing. Whether deficits in joint attention behaviour are specific to ASD could not be clearly answered by high-risks sibling studies (Jones et al., 2014). On the other hand, a systematic review of home-video studies, found consistent evidence that joint attention deficits (initiation or response) in children with ASD were more severe than in children with

DD (Saint-Georges et al., 2010), and this was also confirmed by a more recent home-video study (Watson et al., 2013) not included in the systematic review by Saint-Georges et al. (2010).

Response to name. In the second year of life, the evidence for reduced response to name call is limited to home-video studies, as high-risk sibling studies have not investigated this in the first year; evidence from home-video studies was fairly consistent, with four out of five studies reporting reduced response to name in ASD children compared to TD children (Saint-Georges et al., 2010). But it was not tested whether children with ASD differed in this from children with other DD.

Eye contact. One sign which has been consistently found in home-video studies to differentiated children with ASD not only from typically developing children but also from children with DD in the second year, included lower eye contact quality, and gaze aversion., In particular, eye contact quality was found to be a good predictor of a later diagnosis of ASD (Saint-Georges et al., 2010) and to differentiate children with ASD from those with DD. One high-risk sibling study explored frequency of eye contact and found this to be reduced in children with ASD during the second year (see Jones et al., 2014).

Affect behaviours: Different affect behaviours have been examined but according to the systematic review of high-risk sibling studies by Jones et al. (2014), evidences is limited to investigations by single studies. These found that children with ASD compare to TD children showed more negative affect at 14 months, less affective response to distress at 18 months and less sharing of positive affect at 14 and 24 months (see review by Jones et al., 2014). Evidence from home-video studies, however, found that emotional expressiveness and facial expression was generally reduced in children with ASD compared to TD children, and that reduced positive affect, furthermore, differentiated children with ASD not only from TD children but also from those with DD (Saint-Georges et al., 2010).

Seeking out physical contact. Cuddling behaviour or seeking out physical contact with others can be categorised as a behaviour aiding social relating. In the second year, home-video studies have consistently observed that children with ASD showed reduced seeking out of physical contact and cuddling when compare to children with TD and DD (Saint-Georges et al., 2010). And this is in line with evidence from high-risk sibling studies (Jones et al., 2014). Avoidance of physical contact could also be the result of sensory processing issues, but this is not clear from the existing research.

Social attention and interest. Home-video studies found that children with ASD had diminished interest in their peers compared to TD children and children with DD (Saint-Georges et al., 2010). Preference for aloneness was less consistently reported by home-video studies for children with ASD (Saint-Georges et al., 2010). High-risk sibling studies furthermore have found deficits in imitation behaviours in ASD children compared to TD

children, but this was not compared to children with DD (Jones et al., 2014). Imitation is another way of showing interest in others. Imitation is an important source of learning, especially social skills and it has been hypothesised that the reduced imitation behaviours in children with ASD may contribute to subsequent difficulties with social skills. High-risk sibling studies have found deficits in imitation behaviours in ASD children compared to TD children in the second year, but this was not compared to children with DD (Jones et al., 2014).

Vocalisation and language. Evidence from high-risk sibling studies is mostly consistent for deficits in language development in children with ASD when compared to TD children (Jones et al., 2014; Mitchell et al., 2011), although the type of deficits and delay reported by existing literature reviews has been somewhat inconsistent. One review reported evidence for expressive language abilities being possibly greater than receptive language abilities in children with ASD (Mitchell et al., 2011), while another review reported evidence for the opposite, i.e. receptive language abilities being greater than expressive language abilities (Jones et al., 2014). Also there is some evidence that language deficits may not be an ASD specific sign as these are similar to those of language delayed children without ASD, (Jones et al., 2014). Furthermore there is evidence that delay in language development may be more pronounced in children with more severe ASD (Jones et al., 2014). Home-video studies consistently reported that children with ASD had a lower number of words and sentences than TD children, more stereotyped vocalisations, and reduced ability to follow verbal directions in the second year of life (Saint-Georges et al., 2010).

Play behaviour. Home-video studies consistently found differences in play behaviours in the second year in children with ASD when compared to children with TD and DD, such as less engagement in conventional games (such as peek-a-boo) (Saint-Georges et al., 2010) and a decreased flexibility and variety of play (Palomo et al., 2006). The existing reviews on high-risk sibling studies did not report on evidence for impaired play behaviours.

Table 1.3.2 provides an overview of the deficits consistently observed in children with ASD when compared to children with TD and DD based on the evidence reported by recent systematic literature reviews.

Table 1.3.2: Summary of consistently reported early deficits in social interaction and communication behaviours in infants and toddlers with ASD

ASD vs. TD	ASD vs. DD
First year	
Reduced use and repertoire of gestures	Reduced response to name call
Reduced facial expression	Reduced looking at others
Reduced vocalisation	Reduced frequency and quality of affective
Reduced joint attention behaviours	expression
Second year	
Reduced response to name call	Reduced joint attention behaviours
Reduced imitation behaviours	Reduced eye contact quality
	Gaze aversion / avoidance of eye contact
	Reduced display of positive affect
	Reduced physical contact seeking
	Reduced peer interest
	Deficits in play behaviours

TD = Typical development; DD = Developmental delay

1.2.3   Other early signs of ASD

Repetitive motor actions (stereotypies) have been found in home-video studies to differentiate ASD from typical development but not developmental delay during the first year (Barbaro & Dissanayake, 2009; Saint-Georges et al., 2010), and this has been confirmed by findings from high-risk sibling studies (Jones et al. 2014). In the second year, both home-video studies and high-risk sibling studies again found that repetitive motor behaviours were increased in children with ASD compared to TD, however no comparison with DD children has been conducted (Jones et al., 2014; Saint-Georges et al., 2010). Home-video studies also reported consistently more stereotyped vocalisation in children with ASD in the second year when compared to TD children (Saint-Georges et al., 2014). Interestingly, stereotypies are included in both DSM-IV and DSM-5 criteria as symptoms of restricted, repetitive patterns of behaviours, but the evidence suggests that these behaviours are not clear indicators for ASD during before the second year of life.

Atypical visual object exploration has been found to be ASD specific in high-risk sibling studies (Jones et al., 2014) but not home-video studies; home-video studies found that unusual object use and exploration was more common in children with ASD than TD but equally common in children with DD (Saint-George et al., 2010, Palomo et al., 2006; Mitchell et al., 2011).

Seeking out of physical contact has already been discussed as part of social interaction behaviours; however, as mentioned before, whether the avoidance of physical contact is the result of sensory processing issues is not clear from the existing research. No other unusual sensory behaviours have been investigated as early signs of ASD. The DSM-5 includes sensory hypo- or hyper-reactivity as a symptom for restrictive and repetitive behaviours; however whether this also manifests in the first two year of life remains unknown.

Abnormal fidgety movement and an reduced movement repertoire in both ASD and DD groups during the first six months were reported in one home-video study (Phagava et al., 2008). The same researchers also examined motor symmetry, i.e. how symmetrically aligned the infant's limbs were when they were moving in a supine position (lying on their back). They found that in the first six months children with ASD had reduced symmetry when compared to infants with typical development as well as infants with other developmental delay, suggesting this may be an ASD specific sign (Esposito, Venuti, Maestro, & Muratori, 2009)

Low levels of muscle tonus (hypotonia or floppiness), reduced activity levels (hypo-activity) and delayed motor development in the second year have been investigated by home-video studies, and been found to not differentiate children with ASD form children with DD; these behaviours are, therefore, not specific or unique to ASD but rather are symptoms of developmental problems in general (Saint-Georges et al., 2010; Mitchell et al., 2011).

Summary

In summary impairment in an increased number of social interaction and communication behaviours manifest after the first six months and especially in the second year; also more deficits specific to ASD rather than to developmental delay in general can be observed in the second year. Thus detection of ASD based on these deficits may be easier in the second year of life than in the first year. Evidence for deficits in social communication and interaction behaviours in the first six month is weak, due to the small number of studies investigating this specific time period and due to the inconsistency in findings across these studies. This also suggests that identification of symptoms during the first six months may be less reliable. None of the studies investigating the first six months included a DD comparison group, thus it is unclear if the behavioural differences found at this age are ASD specific or whether they are shared with other DD.

Motor stereotypies appear to be a reliable sign of developmental difficulties but have not been found to be specific to ASD. Deficits in other areas that may indicate the social interaction and communication domain have been less well researched, with mostly only one study investigating a particular behaviour. Therefore, the evidence for behaviours from other domains is not sufficient to suggest that any of the examined behaviours are ASD specific early markers or not.

Evidence from research into early signs of ASD is the basis on which screening instruments are developed. In the following section existing screening instruments for infants and toddlers will be described and their clinical usefulness for early routine ASD screening will be discussed.

1.3 Overview of ASD Screening Instruments

1.3.1 Clinical Utility of Screening Instruments

The growing evidence for early signs of ASD as discussed before has triggered efforts to develop screening instruments for ASD that can be used with infants and toddlers. Furthermore, the availability of evidence based early intervention programs, such as the Early Start Denver Model (ESDM, Dawson et al. (2010)), require that reliable early screening instruments to be available.

Screening for ASD can occur at two levels. Level 1 aims to identify children at risk for any type of atypical development including ASD; this is also referred to as universal or routine screening; these screening tests are not diagnostic instruments but rather flag the need for further assessment. Level 2 screening aims to ascertain the need for a diagnostic assessment for ASD in individuals already identified as at-risk for developmental disabilities. Level 2 screeners therefore have to be more specific than level 1 screeners. And diagnostic assessments typically involve multiple health professionals from different disciplines as well as requiring observational assessments (e.g. Autism Diagnostic Observation Schedule (ADOS; Lord et al. (2000)), and parent questionnaires and diagnostic interviews (e.g. ADI-R; Rutter, LeCouteur, and Lord (2003)); because of their time intensity and complexity assessments are only conducted when concerns about possible ASD have already been raised.

The clinical utility of a screening instrument is usually judged by its sensitivity, specificity and positive predictive values (PPV). Sensitivity is the ratio of positive screens to the total number of cases (a/(a+c), while specificity is the ratio of negative screens to the total number of non-cases (d/(b+d). The positive predictive value (PPV) is the ratio of true positive screens to all positive screens (a/(a+b); the less commonly reported negative predictive value (NPV) is ratio of true negative screens to all negative screens (d/c+d). See Table 1.4.1 for a visualisation of these concepts. PPV can be established by follow-up of positive screens only and thus is often the preferred indicator used in population based studies when follow-up of all negative screens is less feasible.

Table 1.4.1: Screening outcomes

Result of Screening	Diagnostic status		Total
	ASD	No ASD	Total
Positive (at risk)	a true positive	b false positive	a+b Total positive screens
Negative (not at risk)	c false negative	d true negative	c+d Total negative screens
Total	a+c Total ASD	b+d Total non-ASD	

Because level 1 screening instruments aim to detect children at risk of ASD with the aim to refer them for further evaluation, they need maximum sensitivity in order to minimize the number of cases missed (false negative screens). On the other hand their specificity must not be too low either, since they are designed for population wide screening; thus, even a small percentage of false positive screening results would lead to many unnecessary investigations and potentially cause anxiety in parents.

In the case of ASD it has been widely acknowledged that screening should not be a singular event, but rather be part of the ongoing monitoring of children's development with repeated screening to ensure accurate detection of children at risk (Barbaro & Dissanayake, 2010; Barton, Dumont-Mathieu, & Fein, 2012). This is based on the evidence that some signs emerge later or are less stable across time, that some children will show signs only later (late onset), and that the manifestation of some behaviours may be impacted by the clinical context in which the child is examined. Repeat screening also minimizes the potential harm caused by a single incorrect screening result

An overview of currently available level 1 screening instruments for ASD will be presented here with discussion of their clinical usefulness as defined above; screening instruments have been grouped into those specifically developed to detect ASD (ASD specific screeners) and those developed to detect more global developmental problems, but which have also been tested for their use in detecting ASD (Non ASD specific screeners). Only instruments that were designed to assess infants and toddlers have been included in this overview; thus widely used screening questionnaires such as the Childhood Autism Rating Scale (e.g. CARS; Schopler, Van Bourgondien, Wellman, and Love (2010)) have been excluded from this overview because they have not been validated for the use before age two years.

1.3.2 ASD Specific Screeners

The revised Modified Checklist for Autism in Toddlers
The revised Modified Checklist for Autism in Toddlers (M-CHAT) (Eaves, Wingert, & Ho, 2006; Kleinman et al., 2008; Pandey et al., 2008; Robins et al., 2013; Robins & Dumont-Mathieu, 2006; Robins, Fein, Barton, & Green, 2001) is the most widely used and evaluated screening tool for ASD in toddlers. The M-CHAT is a parent report questionnaire designed to assess risk for ASD in children aged 16 and 30 months. The M-CHAT is the modified version of the CHAT (Baron-Cohen, Allen, & Gillberg, 1992) and consists of 23 yes/no questions about the child's current skills and behaviours across five behavioural domains: pretend play, proto-declarative pointing, joint-attention, social interest, and social play. A positive screen is defined as either two failed items or more of the six critical items (2/6), or as any three failed items or more out of the total 23 items (3/23). The critical items include joint attention behaviours (proto-declarative pointing, following a point, and bringing objects to show to parent), behaviours of social relatedness (interest in other children and imitation), and communication behaviours (responding to name).

Several studies have investigated the clinical utility of the M-CHAT as level 1 screening instrument. Clinical utility was established through a later diagnostic assessment of positive screens; this typically involved the Autism Diagnostic Observation Schedule (ADOS, Lord et al. (2000)) and/ or the Autism Diagnostic Interview – Revised (ADI-R., Rutter et al. (2003)), both instruments are considered to be the gold-standard for ASD diagnostic assessment (Levy, Mandell, & Schultz, 2009).

In a pre-selected high-risk sample of toddlers aged 24 to 48 months the M-CHAT was found to have a sensitivity of 0.77 and a PPV of 0.70 for the 2/6 critical item cut-off and a sensitivity of 0.92 and PPV of 0.69 for the 3/23 item cut-off. Specificity was low for both cut-offs with 0.43 and 0.27 respectively (Eaves et al., 2006). Given that the study used a sample of high-risk toddlers the high PPV is unsurprising, but the low specificity is of concern as it results in a large number of false positive screens suggesting the M-CHAT may be picking up other developmental difficulties in addition to ASD.

In an unselected population sample the M-CHAT was found to have a PPV of 0.54 for ASD if a repeat screen was administered to all positive screens, and a PPV of 0.98 for any developmental disorder (Chlebowski, Robins, Barton, & Fein, 2013). A single M-CHAT screen in a community sample has been found to result in a much lower PPV of 0.11, and in a high-risk sample a single screen resulted in a PPV of 0.60 (Kleinman et al., 2008), this improved to a PPV of 0.65 and 0.76 respectively when a follow-up telephone interview was administered to parents of children who had a positive screening result.

A recently revised 20 item version, the M-CHAT-R/F, has been evaluated on a community sample of 15 612 toddlers, aged 16 – 31 months (Robins et al., 2013). The authors

of the M-CHAT-R/F specified three cut-offs: low risk, requiring no further evaluation if two or less items were failed and no other risk factors are present; medium risk if 3–7 items were failed; high risk if 8–20 items were failed, which warrants an immediate referral for diagnostic evaluation and early intervention. In the case of a positive screen (three or more items failed) parents received a structured follow-up interview with the clinician (lasting up to 10 minutes according to the authors). The validation study included an initial screen, a repeat screen if the first was positive, and a diagnostic evaluation between age 21 and 33 months for those children who screened positive at both times. The PPV of two consecutive positive M-CHAT screens with interview follow-up for detecting ASD was 0.51; based on the initial screen alone the PPV was 0.14. Sensitivity of a single screen was 0.91 and specificity was 0.86, and for the two-stage screening with a follow-up interview it was 0.67 and 1.00 respectively.

It is noteworthy that all M-CHAT studies established sensitivity and specificity on diagnostic follow-up of positive screens only; if all negative screens were followed up it would be expected that some would have met a diagnosis of ASD (false negatives). Thus means the true sensitivity and specificity of the M-CHAT is likely to be lower than what has been reported by the authors of these studies.

In conclusion, both, the M-CHAT and the M-CHAT-R/F seem to have a low threshold for indicating developmental risk, which on the one hand results in many children with ASD being detected, but on the other hand it also leads to a high number of false positives. Therefore, it appears that the M-CHAT is screening for developmental difficulties more broadly rather than ASD specifically. The high number of false positive screens is of concern as it may cause unnecessary parental anxiety, especially if it is presented to them as an ASD screener; also the high number of follow-up screens and interviews put stress on limited clinical resources. Therefore, the M-CHAT's usefulness as ASD specific level 1 screener appears to be limited despite its popularity.

The Early Screening for Autism Traits (ESAT)
The ESAT is a clinician administered parent questionnaire. Although it was developed as a level 2 screener, the first four items have been used as level 1 screener as they have been shown to discriminate between children with ASD and those without. The four screening items assess interest in different toys, varied play, emotional expression and reaction to sensory stimulation (Swinkels et al., 2006). A Dutch screening study of 31,724 children used the four ESAT items for screening at routine well-baby checks at around 14 to 15 months of age (Dietz, Swinkels, Daalen, Engeland, & Buitelaar, 2006; Swinkels et al., 2006). If the parents answered negative to at least one of the items it was considered a positive screen, the full ESAT was then administered at a home-visit (1.5hrs per assessment). Positive screens on the full ESAT received a full diagnostic assessment and a diagnostic re-evaluation at 24 and 42 months. Negative screens on the full ESAT were followed up with a parent

questionnaire, and if scoring high on the questionnaire, these were also referred for diagnostic evaluation. The study found that 1.2% of the total sample screened positive on the 4 item-ESAT; however none of the items were able to differentiate between ASD from non-ASD children in this study. The second step screening with the full ESAT had a PPV of 25%. Although easy to administer as part of routine baby-checks, based on the study findings the four-item ESAT does not seem to be clinically useful in the early detection of ASD, nor does the full ESAT due to its low specificity and time intensive administration.

First Year Inventory (FYI)
The FYI is a 63 item parent report measure designed specifically to assess children aged 12 months (Reznick, Baranek, Reavis, Watson, & Crais, 2007). To test its utility for screening the FYI was administered retrospectively to parents of three groups of children: ASD, DD and TD (Watson et al., 2007). PPV was reported to be 0.74 and NPV to be 0.93; sensitivity was 0.92 and specificity was 0.78 for ASD diagnosis in this study. However, the retrospective administration may have been subject to recall bias by parents, and thus the results of this study may be limited in their validity.

A prospective study compared the screening results of the FYI at 12 months with outcomes of a full diagnostic assessment at 3 years (Turner-Brown, Baranek, Reznick, Watson, & Crais, 2012). The full sample (n=699) was assessed at 3 years with the Social Responsiveness Scale (SRS, Constantino and Gruber (2005)) and the Developmental Concerns Questionnaire (DCQ, Reznick, Baranek, and Watson (2005)). Children who scored above the 90th percentile on the FYI at 12 months, or who at 3 years were flagged as at risk on either of these instruments, or had an ASD diagnosis or a family history of ASD received a full diagnostic evaluation with the ADOS. The FYI was found to have a sensitivity of 0.44 and a positive predictive value of 0.14. The low sensitivity, its limitation to screening at 12 months only and its considerable length (63 items) limit the FYI's clinically usefulness as routine screening instrument for ASD. Also children who were not flagged at risk were not evaluated in the validation study, thus true specificity and sensitivity of the scale could be established.

Social Attention and Communication Study (SACS) screener
The SACS study (Barbaro & Dissanayake, 2010, 2013; Barbaro, Ridgway, & Dissanayake, 2011) is a prospective cohort study examining the feasibility of an ASD surveillance program within the Australian public health system. Universal routine screening for ASD was provided by maternal and child health (MCH) nurses at 8, 12, 18 and 24 months well-baby checks in the metropolitan area of Melbourne. A screening instrument was developed for this purpose and nurses were trained in its use prior to implementation. The screener consists of items assessing social attention and communication behaviours, based on the evidence for early signs of ASDs in infants and toddlers. The screening items vary in type and number depending on assessment age (7 to 16 items), with some of them classified as

key behaviours by the authors. The absence of these key behaviours is considered a risk flag; however from these reports it remained unclear how many key behaviours need to be absent to indicate a positive screen. Based on the positive screens, the study found a PPV of 0.90 at 12 months, 0.79 at 18 months, and 0.81 at 24 months. Because negative screens could not feasibly be followed up, sensitivity and specificity was estimated based on a prevalence rate 0.01. Based on the children that received a diagnostic assessment the sensitivity of the SACS screener was moderate with 69% and its specificity was high with 99.9%, respectively across the whole age period of 12-24 months (Barbaro & Dissanayake, 2010). Sensitivity and specificity for each age were not reported, thus its clinical usefulness for screening at age 12 months remains unclear. Importantly, the study found the acceptability of the SACS screener among MCH nurses to be high. However, the SACS screener remains a complex tool to use requiring assessment of critical signs of different types and numbers for each of the four screening ages. Because the screener is modelled along the four routine well-baby checks offered by the public health system at ages 8, 12, 18 and 24 months the utility of this screener outside these ages is unclear and the positive screen criteria for each age also remains unclear. Nevertheless, the SACS appears to be the most reliable observational ASD specific screening instrument currently available.

1.3.3 Non ASD specific screeners

There are several developmental broadband screeners designed to screen for developmental difficulties in general. Some of these screeners have also been examined for their ability to detect ASD and are described in the following.

The Infant-Toddler Checklist (ITC)

The Infant-Toddler Checklist (Wetherby et al., 2004) is one component of the Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP; Wetherby and Prizant (2002)) which is designed to detect any communication delays in children aged 6-24 months, and thus is not specific to ASD. The ITC component is a 24 item parent checklist assessing developmental milestones, social communication, and parental concerns. It has been validated for the detection of ASD on a general population sample, and a sample with communication delay at aged 9 to 24 months (Wetherby, Brosnan-Maddox, Peace, & Newton, 2008). Communication assessments were conducted for positive screened children, and a comprehensive ASD assessment was conducted at 3 years for children with a positive ITC screen or reported parent concern or who were flagged as at risk in the communication assessment. A positive screen was defined as either a total ITC score or a Social or Symbolic Composite score below the bottom 10th percentile, or a Speech composite score below the bottom 10th percentile on two consecutive ITC screens. The ITC was found to have a high false negative rate when used at 6–8 months of age and only 20% of those later diagnosed with ASD screened positive at that age (PPV = 0.2). The ITC performed better at 12–14

months through to 21–24 months in detecting any communication delay including ASD, but was not able to differentiate between children with ASD and those with other communication delay. Of the children who were subsequently diagnosed with ASD, 93.3% had a positive ITC screen at some stage before age 2. However, 18% of the total sample of 5,385 children had a positive screen on the ITC but only 60 children (or 1%) later met diagnostic criteria for ASD; this equals a PPV of 0.06. to the high false positive rate was due to the scale picking up on other communication delays. Specificity of the ICT was not reported by the authors of the validation study, but they noted that the proportion of positive screens was unusually high in their sample, inferring a low specificity. Based on these findings the ITC appears to be a useful Level 1 screener for communicative problems in general but not for ASD specifically.

Brief Infant-Toddler Social and Emotional Assessment (BITSEA)

The BITSEA is a parent-report measure designed to screen for any developmental problems of children aged 12–36 months (Briggs-Gowan, Carter, Irwin, Wachtel, & Cicchetti, 2004); it has two subscales, a problem behaviour scale (31 items, including 9 ASD-specific items), and a competence scale (11 items, including 8 ASD-specific items). Parents respond to each item with 'Not True/Rarely', 'Somewhat True/Sometimes' or 'Very True/Often'. The BITSEA's usefulness as screener for ASD was examined as part of a longitudinal study (Gardner et al., 2013). Children were assessed with the BITSEA at 12 and 24 months and with the M-CHAT at 24 months; the M-CHAT was used in lieu of a diagnostic evaluation. The authors found that the ASD specific items from both subscales predicted M-CHAT risk status at 24 months. Using ROC analysis they established sensitivity of the BITSEA ASD scales at 12 months as 0.71 and at 24 months as 0.83 for M-CHAT risk, however the cut-off scores for each age were not reported. PPV was not reported either. Because the M-CHAT is not a diagnostic tool but another parent report risk screener, the BITSEA's true sensitivity and specificity for ASD diagnosis is likely to be much lower, given the M-CHAT's low PPV as reported before. Therefore, the BITSEA's true clinical usefulness as ASD risks screener remains unknown. Gardner et al. (2013) concluded that the BITSEA is suitable as a Level 2 screener for ASD and as a Level 1 screener for social emotional problems in general.

Parents' Evaluation of Developmental Status (PEDS)

The PEDS is a 10 item parental questionnaire designed as broad developmental screener for children from birth to age 8 years; the PEDS is widely used in Australia. It assesses parents' concerns in the domains of language and motor development, self-help skills, early academic skills, behaviour and social-emotional/mental health. Parents respond to items with 'no', 'a little' and 'yes' to indicate their level of concern (Glascoe, 2006).

Several studies have examined the PEDS' utility in detecting children at risk of ASD, using the M-CHAT for ascertainment of cases; as described before, the M-CHAT is not a diagnostic assessment instrument but a risk screener, and thus using it to establish

sensitivity and specificity for ASD screeners is rather problematic. Thus findings from these studies can only be used to establish concurrent validity of the PEDS and the M-CHAT.

In an online survey children aged 18 to 59 months were screened with the PEDS and the M-CHAT (Glascoe, Macias, Wegner, & Robertshaw, 2007). It is unclear whether the researchers administered a follow-up interview for the M-CHAT as suggested as best practice by findings from earlier research. The analysis only included children who scored at moderate to high risk on the PEDS, thus negative screens were not analysed to determine sensitivity of the PEDS. Of the positive PEDS screens 66% had also positives screens on the M-CHAT (PPV=0.66). The authors identified an ASD specific subscale score including concerns in the domain of behaviour, fine motor, gross motor, receptive language and social-emotional skill. In the age group 18 to 35 months (n = 233) using a cut-off of three or more of concerns rendered a sensitivity of 0.79 and a specificity of 0.75 (ascertainment with the M-CHAT).

These results were contradicted by a study using a sample of 152 of 18-30 months old children presenting to a routine baby health check at their physician (Pinto-Martin et al., 2008). The PPV for the PEDS was found to be 16% for positive screen on the M-CHAT, with a false negative rate of 14%. The authors found no correlations between endorsing concerns on the PEDS and failing any of the M-CHAT critical items. The authors of the PEDS criticised this study for using an old PEDS scoring algorithm (Glascoe & Squires, 2009).

A recent Australian study examined the utility of the PEDS for universal ASD screening in 87 children aged 16 to 60 months (Eapen, Črnčec, Woolfenden, & Blackmore, 2013) including low and high-risk children. Positive screens on the PEDS were compared with positive screens on the M-CHAT. Positive screens on the M-CHAT were defined as six failed items or more which differs from the cut-off suggested by previous research (Eaves et al., 2006; Kleinman et al., 2008). The overall sensitivity for the PEDS to detect positive M-CHAT screens was 0.75. The authors derived adjusted scores by multiplying the ROC values for the M-CHAT with those for the PEDS. In the 16 to 30 month old cohort the adjusted sensitivity was 0.87, adjusted specificity was 0.94, adjusted PPV was 0.40, and adjusted NPV was 0.99. However, the sample size for this age group was small with n=20. Because no diagnostic ascertainment of children identified as risk with the PEDS was conducted in any of these studies, the actually specificity and sensitivity of the PEDS as a screener for ASD is likely to be lower. However, the PEDS seems to have satisfactory concordance with the M-CHAT.

1.3.4 Summary

In summary, a range of level 1 screening instruments have been developed and tested regarding their clinical utility for detecting ASD in infants and toddlers; some were specific to signs of ASD, others were designed to screen for a range of developmental problems. Several of the validation studies had methodological limitations and thus sensitivity and

specificity of the examined screening instruments could not be reliably established. Some studies used pre-selected high-risk samples and thus could not establish sensitivity of the instrument for use in the general community (e.g. Eaves et al., 2006). Other studies used a second screening instrument instead to ascertain diagnostic status of positive screens instead of a gold-standard diagnostic evaluation; this means the actual sensitivity and specificity of these screening instruments (i.e. PEDS, BITSEA) for detection of ASD remains unknown.

All but one of the reviewed screeners were parent report measures. The advantage of parent report measures is that they are easy to implement and least time consuming for clinicians which may explain their appeal. Only two parent questionnaires were identified that have been validated for screening children aged 12 months or younger (FYI, ITC), one of which showed high sensitivity in detecting ASD at this age (ITC). Parent questionnaires performed better in the second year of life with high sensitivity for two screeners (M-CHAT-R/F, ITC); however, the fairly high sensitivity of the M-CHAT was offset by a fairly low specificity; and the authors of the ITC validation study did not report specificity but indicated that it did not differentiate between children with ASD and those with other communication delays. Thus both, the ITC and the M-CHAT may have only limited utility as level 1 screener for ASD.

The SACS screener was the only observational screening instrument identified. It showed satisfactory sensitivity and specificity estimates. This is also the only screener whose acceptability to clinicians was examined and found to be high; acceptability is an important aspect for any screening instrument designed for routine use as even the most sensitive and specific screener is unlikely to be used if clinicians find them cumbersome or too time consuming to implement. While the SACS screener seems to be a good candidate for a level 1 screener, it is a fairly complex screening instrument with changing number and types of items to be observed at each age, and cut-off criteria at each age have not been specified clearly.

In conclusion, almost all existing screening instruments rely on parent report and are limited in their clinical utility for routine ASD screening in the first two years of life by their low specificity. Only the observational SACS screener is currently showing promise. Thus there is a need to improve existing screening tools, explore the clinical utility of other developmental screener, or to develop new ones.

Table 1.4.4: Clinical utility of existing level 1 screeners for ASD in infants and toddlers

Name	Type	# Items	Age of validation sample (months)	Criteria for positive screen	PPV	Sensitivity	Specificity
ASD specific screeners							
M-CHAT	Parent questionnaire	23	16-31	≥ 3 failed items (3/23) or ≥ 2 failed ASD items (2/6)	0.11 M-CHAT only 0.65 with follow-up parent interview	0.77 critical items cut-off$^\$$ 0.92 overall$^\$$	0.27 for 3/23 0.43 for 2/6
M-CHAT-R/F	Parent questionnaire	20	18-24	≥ 3 failed items	0.14 for single screen 0.51 for 2-stage screen plus interview	Single:0.91 2-stage:0.86	Single: 0.67 2-stage:1.00
ESAT – 4	Parent questionnaire	4	14 -15	≥ 1 'yes' response	0.25	Not reported	Not reported
FYI	Parent questionnaire	63	12	Total score > 96th percentile	0.14	0.44	0.99
SACS screener	Observational scale	7 – 13	12 - 24	Absence of key behaviours at each age, but no cut-off reported	0.90 at 12 months 0.79 at 18 months 0.81 at 24 months	Overall 0.69	Overall 1.00
Developmental broadband screeners							
ITC	Parent questionnaire	24	6 - 24	Social/ symbolic composite/ total score ≤10th percentile Or: Speech composite ≤10th percentile at two consecutive screens	Overall: 0.06	6-8m: 0.20 9-11m: 0.77 12-14m: 0.91 15-17m: 1.00 18-20m: 0.92 21-24m: 0.95 **Overall: 0.93**	Not reported
BITSEA	Parent questionnaire	42	12 ,24	Not reported	Not reported	12m: 0.71[†] 24m: 0.83[†]	Not reported
PEDS	Parent questionnaire	10	16-30	≥ 3	0.40[†]	0.87[*]	0.94[*]

AD: Autistic Disorder; ID: Intellectual Disability; TD: Typical Development; $^\$$ for high-risk children; [†] M-CHAT used instead of diagnostic evaluation; [*] for ages 16-30 months

1.4 Screening for social withdrawal in infants as opportunity to detect ASD

Given the dearth of validated and clinically useful level 1 screeners for ASD in infants and toddlers, it is worthwhile to investigate an instrument that is being used internationally as screener for more global 'social withdrawal' in infants, the Alarme Distress de Bébé Scale (ADBB, Guedeney and Fermanian (2001)). In normally developing infants social interaction skills emerge during the first two months of life; this includes the ability to initiate and maintain eye contact with another person, to vocalise, and use facial expressions and body and head movements to engage the caregiver or others in interactions (Trevarthen & Aitken, 2001). Infant social withdrawal is characterized by deficits in these areas such as reduced or lack of eye contact, smiling, cooing or crying and fussing behaviours. Social withdrawal thus refers to a reduced or lack of responsiveness to social stimuli; it is important to note, that social withdrawal is different from a shy or inhibited temperament; shy children take longer to "warm up" with an adult but are nonetheless responsive to their bids. Indeed research has shown that social withdrawal and temperament are independent from each other (see Guedeney, Matthey, and Puura (2013). Thus social withdrawal is defined as decreased responsiveness or lack of responsiveness to social stimuli (Guedeney et al., 2013), and its symptoms have similarities to early signs of ASD.

Existing observational assessment scales of infant social withdrawal require either special equipment or cannot be conducted in-vivo due to the large number of items that need to be coded (See Matthey, ČrnČec, Hales, and Guedeney (2013)); therefore they do not lend themselves to routine screening in clinical settings such as early childhood clinics. One validated instrument, however, that has been proven to be easily adaptable to clinical settings is a scale developed in France, the Alarme Distress de Bébé Scale (ADBB) by Guedeney and Fermanian (2001). The ADBB is an 8-item behavioural checklist consisting designed for the use by nurses and paediatricians in the context of routine physical assessments of the baby from 2 – 24 months of age. Each behaviour is assessed through observation and scored on a 5-point Likert scale; the total score range is 0-32 with higher scores indicating more social withdrawal; a score of 5 or above indicates clinically significant social withdrawal Guedeney and Fermanian Guedeney and Fermanian (2001). The scale has been used across nine countries in 13 controlled studies of infant social withdrawal (Guedeney et al., 2013).

The ADBB has good internal consistency, good inter-rater reliability, and concurrent validity with expert clinical judgement and a 17 – item checklist for developmental risk. The ADBB was found to have as sensitivity of 0.82 and specificity of 0.78 for detection of developmental difficulties (Guedeney & Fermanian, 2001). These findings have been replicated in a Brazilian validation study. When using the criteria of the Diagnostic

Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood (DC: 0-3, Zero to Three (1994) as gold standard, validation studies in Italy and Argentina reported sensitivity to be slightly lower but with a similar specificity (as reported by Guedeney et al. (2013)). These findings suggest that the ADBB is valid across different cultural settings.

The prevalence of social withdrawal has been found to range from 3% in a Finnish study using repeat screenings (Puura et al., 2010) to 13% in a French study using a single screening (Guedeney, Foucault, Bougen, Larroque, & Mentré, 2008). This highlights the importance of conducting a repeat screen to increase reliability of screening results.

Research studies using the ADBB found that a range of factors in the perinatal period can lead to sustained social withdrawal in infants. These factors included parental mental health problems, specifically postnatal depression in mothers (Matthey, Guedeney, Starakis, & Barnett, 2005), foetal alcohol exposure, low birth weight, infant feeding and settling problems, adoption, parental separation , and low maternal sensitivity (as reported by Guedeney et al. (2013)). This is important to consider when using this scale, as social withdrawal may not be indicative of ASD only, but rather flags the need for further investigation of the underlying causes for the social withdrawal behaviour. The ADBB has also been shown to reliably detect poor interaction skills in infants aged 8 to 11 week (Puura, Guedeney, Mäntymaa, & Tamminen, 2007); this is of particular interest to the detection of ASD, as deficits in interaction skills manifested in impaired social and communication behaviours have been widely reported as early markers for ASD. Indeed, a small retrospective home-video study used the ADBB to detect early markers of ASD (Wendland et al., 2010). It found on average higher scores indicating more withdrawn behaviour at 6 and 12 months in children with ASD compared to children with TD. However, case by case analysis of the study sample showed that only half of the ASD cases scored above the clinical cut-off for social withdrawal on the ADBB at any of the assessment time points. But all typically developing children scored below the cut-off point for social withdrawal. These findings show promise for social withdrawal behaviour to be potentially indicative of ASD.

A modified version of the ADBB is available for the Australian context and has been adapted to further fit within the routine assessment at early childhood clinics in NSW (modified Alarm Distress Baby Scale: m-ADBB; Matthey et al. (2013)). The m-ADBB assesses five behaviours instead of the original eight; the three items for which it was difficult to achieve satisfactory inter-rater reliability or those that were highly inter correlated with other items were removed from the modified version. The five remaining behaviours are: facial expression, eye contact, vocalisation to others, activity, and engaging in relationship with others. Each item is rated as 'satisfactory', a 'possible problem' or 'definite problem', with clear descriptors for each scoring option. The m- ADBB has been found to show good concordance with the full ADBB scale (Matthey et al., 2013). Two possible problems or one

definite problem have been found to be the equivalent to the cut-off score of 5 of the original ADBB scale.

The m-ADBB has the potential to be clinically useful in detection of a range of developmental difficulties, as it is a very brief observational instrument that can be easily integrated into routine clinical practice. Thus there is merit in investigating whether the m-ADBB could also reliably detect infants and toddlers at risk for ASD.

1.5 Summary and Conclusion

ASD is a neurodevelopmental disorder of unknown aetiology causing significant impairment of social and communicative functioning from early development onwards. While symptoms of ASD are considered to be present from birth, their onset, type and intensity may vary considerably. Despite this heterogeneity of the ASD symptoms and onset patterns, there is mounting evidence from numerous retrospective and prospective studies that behavioural signs are present during the first two years of life. Among these research efforts, retrospective home-video studies have proven to be especially useful in detection of early behavioural markers.

The evidence from prospective and retrospective studies shows that signs are observable most reliably after the first 6 months, with an increase in number and intensity during the second year. Children with late onset ASD showed early markers at two years latest. Deficits in social and communicative behaviours have been the predominant signs identified by most studies and across ages.

Despite the substantial evidence for early observable ASD signs, routine screening of infants and toddlers is not being implemented due to the lack of appropriate screening instruments. Detection currently relies on parent concern or ASD specific knowledge of clinicians. However, efforts have been made to develop ASD specific screeners for infants and toddlers, as well as to determine the clinical utility of broader developmental screeners for detection of ASD risk. Almost all of these screening instruments rely on parent report, and most are limited by unsatisfactory or unknown sensitivity or specificity; acceptability of these scales to clinicians has often not been examined.

Given the lack of appropriate ASD screening instruments for infants and toddlers, it is worthwhile to examine the clinical utility of a screener for sustained social withdrawal in infants and toddlers. Social withdrawal behaviours can manifest as early as age two or three months, and share similarities to early behavioural signs observed for ASD.

1.6   Study Aim

The current study focusses on the role of infant social withdrawal, as assessed by the modified Alarm Distress Baby scale (m–ADBB), as marker for ASD in infants and toddlers. It aims to examine the clinical usefulness of the m-ADBB in detecting signs of autism in infants and toddlers, using home-video analysis.

The research questions that this study aims to answer in regard to the m-ADBB's utility as ASD screener are:

1. Do infants and toddlers with ASD show more social withdrawn behaviour than children with typical development as assessed with the m-ADBB?
2. Is the detection rate of the m-ADBB clinically useful for level 1 screening?
3. Do infants and toddlers with ASD show a distinct pattern of problems across the behaviours assessed by the m-ADBB?

It is hypothesised that children with a diagnosis of ASD will show more possible or definite problems on the m-ADBB in their pre-diagnostic home videos at ages 12 and 24 month, resulting in statistically significantly higher scores on the m-ADBB when compared to scores for typically developing children (TD) in home-videos at matched ages.

CHAPTER 2: Systematic Literature Review of Home Video Studies

2.1 Background and Aim

This chapter examines in detail the evidence for early signs of ASD that has been generated by retrospective home-video studies. Home-video studies analyse early signs of ASD in non-standardised settings, thus the evidence from these studies is relevant to the development of early routine screening instruments that can be used in a non-standardised clinic setting.

Several reviews of the evidence from home-video studies have been conducted in the past; however these were of varying scope. The most recent review by Costanzo et al 2015 synthesises findings from home-video studies, prospective and parent report studies with a focus on advantages and disadvantages of the different study methodologies rather than synthesising the findings for each study type. Similarly, Mitchell et al. (2011) reviewed evidence across the different study types, which does not allow for a clear reading of the evidence attributable to home-video studies alone. Saint-Georges et al. (2010) conducted a systematic review of home video studies, including all publications until January 2008, and compared findings to those of prospective studies. An earlier review by Palomo et al. (2006) of home-video studies only is most relevant in scope to the current study; however it included only studies published until September 2005; however both home-video reviews by Saint-Georges et al. (2010) and by Palomo et al. (2006) are out-dated. Furthermore, these systematic reviews have excluded studies published in a language other than English, and thus have omitted studies from two prolific research groups in France. Given these limitations of the existing reviews, an up-dated systematic review of home-video studies has been conducted and is described in this chapter.

The aim of the current systematic review is to examine the evidence from home-video studies for behavioural signs indicative of ASD in the first and second years of life. A secondary aim is to explore how different methodological approaches to home-video analysis perform in detection of early signs. The two different approaches that have been employed by home-video studies for assessing infant and toddler behaviours observed in video-recordings are: (1) detailed time-sampled observation and coding of behaviours, and (2) a more global assessment of behaviours with help of a clinical ratings scale. Time sampled coding (TSC) is the more thorough methodological approach and is based on rating

the frequency and intensity of specific behaviours within a set time-interval, e.g. 1 minute. The clinical rating scale approach (CRS) on the other hand, uses the overall impression of the observed behaviour to allocate scores across different items on a validated scale. Thus these studies are of most relevance to the development and testing of routine screeners.

2.2 Methods

2.2.1 Search strategy

Using the Medline, PsycINFO, Embase, CINAHL and Scopus databases, all literature published until April 2014 was searched with the following search terms: (autism or ASD or autistic or Asperger* or pervasive developmental disorder or PDD-NOS or subject headings related to autism and ASD) and (video* or movie* or film* or recording*) and (child* or infant* or baby or babies, or neonat* or relevant subject headings). See Table 2.1 for detailed search protocols for each database. The search results were continuously updated through database search alerts until 30 June 2016.

After removal of duplicates, a total of 2030 references were retrieved; titles and abstracts were screened and 59 references were retained for full text screening. To be considered for inclusion in the review, studies had to meet following criteria: use of retrospective home-video analysis, presence of a control group of typically developing children, age at time of home-video recording between 0 and 24 months, use of observational methods, published in a peer-reviewed journal in English or French language. French language publications were included because of the prolific research output of teams in Tours and Paris since the early 1990ies. The author of this thesis understands French and thus translation of the publications was not necessary.

Based on full text screening, 26 references were excluded for the following reasons: absence of a control group of typical developing children (Adrien et al., 1991; Bernabei, Camaigni, & Levi, 1998; Gernsbacher, Sauer, Geye, Schweigert, & Goldsmith, 2008; Maestro, Casella, Milone, Muratori, & Palacio-Espasa, 1999; Maestro et al., 2005b; Poon, Watson, Baranek, & Poe, 2012; Receveur et al., 2005; Thorsen et al., 2008) or no comparison was conducted with the control group (Goldberg, Thorsen, Osann, & Spence, 2008) published in Italian language (Esposito & Venuti, 2008; Maestro, Milone, Muratori, & Casella, 1999) videos recorded in experimental settings instead of home videos (Adrien et al., 1992a; Wetherby, Watt, Morgan, & Shumway, 2007; Wetherby et al., 2004) automated assessment of behaviours, i.e. crying (Esposito & Venuti, 2010, 2010a); and observational ratings of home-videos conducted by parents (Esposito, Venuti, & Bornstein, 2011); furthermore nine dissertations were excluded (Baranek, 1997; Bayrami, 2011; Book, 2010; Chin, 2009; Esposito et al., 2011; Gerwing, 2009; Mitchell, 2015; Osterling, 1998; Poon, 2005). A hand search of the references cited in the included publications was conducted, and one further publication, which satisfied inclusion criteria, was identified (Brisson, Warreyn, Serres, Foussier, &

Adrien-Louis, 2011b). Of the 35 publications which were included, nine related to studies using clinical rating scales, and 26 to studies using a time-sampled coding approach. Some of the included publications related to the same study samples and constituted reports on different aspects of the analyses. Thus these publications were counted as one study for this review. For CRS studies the following publications referred to the same study: Adrien et al. (1991) and Adrien et al. (1993); Adrien, Gattegno, Streri, Reynaud, and Barthelemy (2005) and Gattegno, Reynaud, Streri, Barthelemy, and Adrien (2005). For TCS studies the following publications referred to the same study: Baranek (1999)and Baranek et al. (2005); Bernard, Adrien, Roux, and Barthelemy (2005) and Bernard, Adrien, Roux, and Barthelemy (2006). Therefore, the eight identified CRS publications related to seven studies, and the 26 TSC publications relate to 24 studies.
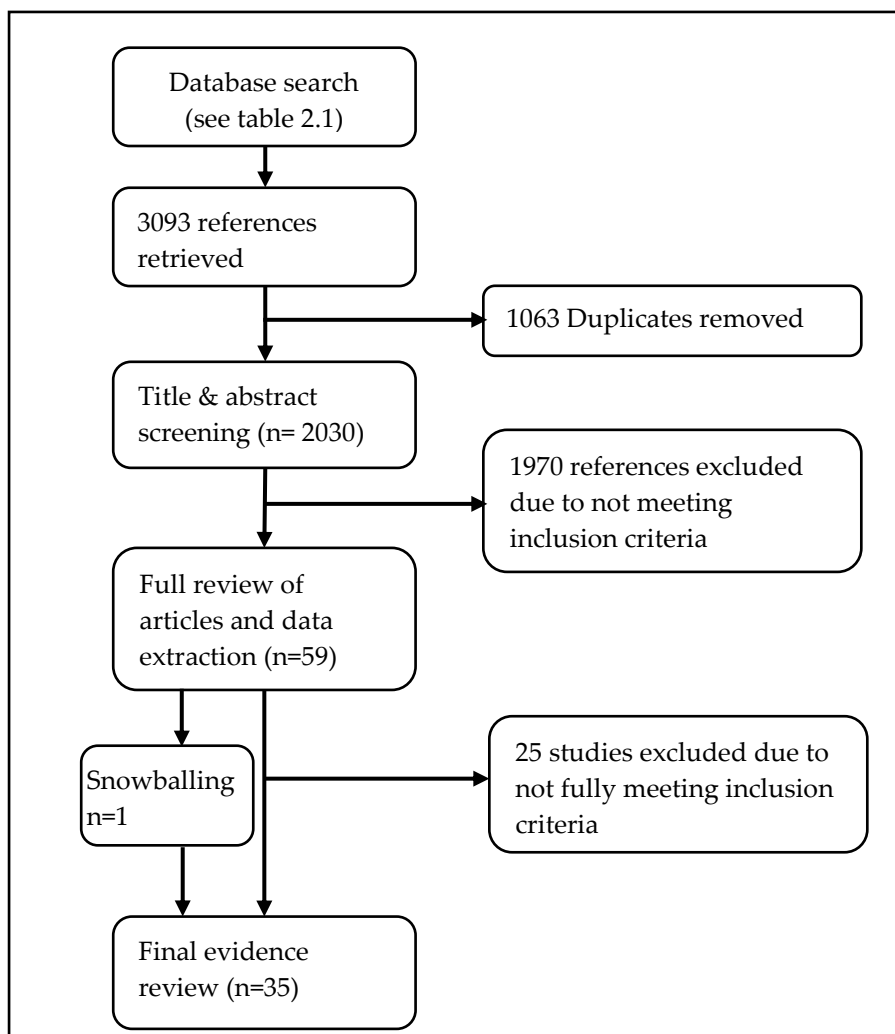
Figure 2.2.1: Study Selection

Table 2.1: Documented search protocol for home-video studies examining early signs of autism

Database	Search terms
Medline	child development disorders, pervasive[SH] OR Asperger syndrome[SH] OR autistic disorder[SH] OR autism OR autistic OR Asperger OR pervasive developmental disorder OR PDD-NOS AND Video Recording[SH/Methods] OR video* OR movie* OR film OR video recording AND Child, Preschool[SH] OR child[SH] OR infant[SH] or infant, newborn[SH] OR baby OR child* OR infant* OR neonat*
PychINFO	Autism[SH] OR Aspergers Syndrome[SH] OR Pervasive Developmental Disorders[SH] OR Autism OR ASD OR Autism spectrum OR pervasive developmental disorder OR PDD-NOS AND video* OR Films[SH] OR movie* OR 'video clip* OR 'video recording' AND child* OR infant* OR baby OR babies OR neonat*
CINAHL	Autistic Disorder[SH] OR Pervasive Developmental Disorder-Not Otherwise Specified[SH] OR autism and developmental disorder [SH] OR Child Development Disorders, Pervasive[SH] OR Asperger Syndrome[SH] OR autistic OR Asperger Syndrome OR Autism Spectrum Disorder OR AND Videorecording[SH] OR video* OR film* OR movie* OR recording* AND infant* OR Infant, Newborn [SH] OR child* OR Child, Preschool [SH] OR baby OR babies OR neonat*
Embase	pervasive developmental disorder[SH] OR 'pervasive developmental disorder' OR PDD-NOS[SH] OR PDD-NOS OR autism spectrum disorder[SH] OR 'autism spectrum disorder' OR autism[SH] OR autism OR aspergers syndrome [SH] OR 'aspergers syndrome' OR autistic OR Asperger* AND Videorecording [SH] OR movie* OR film* OR video* AND Child [SH] OR child OR child* OR infant[SH] or infant OR toddler* OR neonat*
Scopus	autism OR 'autism spectrum' OR 'pervasive developmental disorder' OR asperger* OR autistic OR pdd OR asd AND video* OR movie* OR film* OR recording* AND child* OR infant* OR toddler* OR baby OR babies OR neonat*

SH: Subject Heading; * signifies truncation – if not otherwise specified all search terms are text keywords.

2.2.2 Data extraction

Data was extracted on research design, sample information, type and characteristics of rating scale used, inter-rater reliability, video length and content, behaviours assessed and main outcome pertaining to group differences between cases and controls. Each study was also assessed for its methodological quality according to the following criteria, which were similar to the quality rating by Saint-Georges et al. (2010): a) validation of scale or coding approach; b) blindness of coders to diagnostic status of child; c) report of inter-rater reliability; d) use of statistical analysis with report of effect sizes; e) control for video content and/or length; e) ascertainment of diagnostic status through diagnostic assessment; f) sample size ≥ 10 for each group (cases / controls). Studies could obtain a score of 2 per criterion up to a total of 14 points. See Appendix A for the full checklist.

2.2.3 Data synthesis

The 35 studies included in this review focussed on a range of different behavioural signs and also differed on the level of data analysed, such as comparing cases and controls on individual behaviours, on composites or subscale scores or total summary scores. In order to synthesise study findings the different behaviours assessed by the different studies were summarised into the categories of social and communication behaviours and other signs, including repetitive and stereotyped behaviours as well as motor abnormalities.

2.3 Results

2.3.1 Description of reviewed studies

The reviewed studies have been grouped into those using the CRS approach and those using the TSC approach. See Tables 2.3.1 and 2.3.2 for details of all included CRS and TSC studies respectively.

The seven CSR studies included in this review were conducted by two research groups from France and one from the USA; four papers were published in French and two in English. Of the TSC studies, nine were from the USA, five from Italy, four from France, one from Germany, one from Georgia, and one was conducted as an international collaboration. Four studies were published in French (Bernard et al., 2005, 2006; Brisson, Serres, & Adrien, 2011a; Degenne, Serres, Gattegno, & Adrien, 2009; Guidetti, Turquois, Adrien, Barthelemy, & Bernard, 2004) the other 20 in English.

Methodological quality of the reviewed studies

All studies had a case-control design, either comparing children with ASD to typically developing children only or additionally to children with other developmental delay. The methodological quality scores of CRS and TSC studies were similar, with CRS studies scoring an average of 8.1 (range 5 to 12) And TSC studies scoring an average of 8.6 (range 4-11) out of a possible maximum quality score of 14.

Sample size. All but one of the seven CRS study had samples sizes of less than 15 per group. The TSC studies tended to have larger samples, with 16 of the 23 studies having at least 15 participants per group. Sample size calculations were not presented by any of the studies, suggesting the use of convenience samples. Furthermore, response rates were only described in one study (Baranek, 1999). Thus for most studies it remains unknown what the response rate was and how selective the sample might have been.

Blindness and reliability of raters. Raters were blinded to the diagnostic status of the children in all CRS studies except the study by Adrien et al. (1992b) which used informed coders. This may have inflated the difference observed between the ASD and comparison group; as shown by another home-video study (Zakian, Malvy, Desombre, Roux, & Lenoir, 2000), which explicitly examined the difference between coding of videos by blind versus informed raters, raters with knowledge of the diagnostic status substantially over-reported problems in children with ASD. Of the TSC studies four did not report on whether the raters were blind to the diagnostic status of the children observed in the videos; the two TSC studies by Clifford used one blind and one informed rater. Thus observations may be of limited validity in studies that used informed rated or did not report on whether raters were informed of children's diagnostic status or not.

Inter-rater reliability was reported in four CRS studies. Gattegno et al. (2005) conducted a test-retest reliability analysis only but no IRR analysis; the study by Adrien et al. (1992b) did not report any reliability testing. Five of the TSC studies did not report inter-rater reliabilities (Bernard et al., 2005, 2006; Brisson et al., 2011a; Degenne et al., 2009; Maestro et al., 2001; Maestro et al., 2006). Where no IRR was reported the ability to judge the reliability and validity of the video observations is limited.

Video content and length. Video content was controlled for in four CRS studies (Adrien et al., 1993; Malvy, Adrien, Roux, Lataste, & Sauvage, 1994; Ozonoff et al., 2008b; Zakian et al., 2000); however length of video observation was only controlled for in one CRS study (Ozonoff et al., 2008b). The length of home-videos in CRS studies ranged from 10 minutes to 138 minutes, but was not reported in four studies. Most TSC studies divided frequency of occurrence of a target behaviour by the length of the video recording or the number of scenes coded, thus controlling for the variability in video length; other studies selected segments of the videos of equal duration, i.e. 5 minutes per age. Most studies also coded for video content and number of people present to examine comparability of videos between study groups. Video length ranged from half a minute to 90 minutes in TSC studies, but was not reported in another nine studies. For studies which did not report video length and content it remains unclear how video observations may have been impacted by these video characteristics.

Diagnostic Ascertainment. Six CRS studies ascertained the diagnostic status of cases through independent diagnostic assessments; one study did not report the method of ascertainment

(Malvy et al., 1994). Cases constituted children later diagnosed with Autistic Disorder (according to DSM-III or IV) in five studies; Wendland et al. (2010) also included children with PPDNOS and Asperger's Syndrome in their sample, and Adrien et al. (1992b) included children with AD and PDD-NOS.

Ten of the 23 TSC studies included only children with Autistic Disorder (according to DSM-II or IV criteria), six studies also included children with a diagnosis of PDD-NOS, four studies did not clearly report on the type of diagnosis included in the case group; the most recent study included children who met a DSM-5 ASD diagnosis (Zappella et al., 2015). Thus evidence from TSC studies mostly pertains to early signs of autistic disorder rather than autism spectrum disorder. In particular, children with Asperger's Syndrome have been excluded from all but one study.

Data analysis. Data analysis in six CRS studies was limited to group comparisons and comparisons between time points. Two studies (Adrien et al., 1992b; Gattegno et al., 2005) did not test for statistical significance of group differences and limited their analysis to descriptive statistics. Detection rates were not reported in CRS studies. One CRS study also investigated developmental trajectories by using growth curve modelling (Ozonoff et al., 2008b). This study had also the by far the largest sample of 48 children with ASD and 24 children with TD and 25 children with DD, allowing for more sophisticated statistical analysis. All but one TSC study (Degenne et al., 2009) reported on statistical group differences; six TSC studies also conducted discrimination analyses and reported numbers correctly classified. Effect sizes of observed group differences were rarely reported, with notable exceptions for the CSR study by Ozonoff et al. (2008b) and the TSC study by Clifford, Young, and Williamson (2007).

2.3.2   Evidence from studies using clinical ratings scales (CRS)

The seven CRS studies used five different clinical rating scales, of which only two were validated. The scales were developed by the respective research groups themselves.

Infant Behaviour Summarized Evaluation (IBSE) scale: Four studies (Adrien et al., 1993; Adrien et al., 1992b; Wendland et al., 2010; Zakian et al., 2000) used the IBSE scale (Adrien et al., 1992a), a 33-item scale that assesses infant behaviours across the domains of socialisation, communication, adaptation to the environment, tact-tonus motility, emotional and instinctual reactions, attention and perception. Each behaviour is rated on a 5-point Likert scale for the frequency of its occurrence. The scale was validated based on observations of standardised interactions of a clinical sample of 6 to 48 months old children either referred for developmental problems, or typically developing but with settling or feeding problems (Adrien et al., 1992a; Wendland et al., 2010; Zakian et al., 2000). From the 33 items, a 19 item factor for 'Autism' was established a which classified 83.3% of children with ASD (based on DSM-III criteria) and children with developmental delay correctly with a sensitivity of 84.6% and specificity of 81.8%. When comparing children with ASD to typically developing

children the scale classified 94.6% correctly with a sensitivity of 92.3% and specificity of 100%.

Across the three studies using the IBSE, three items have been consistently identified as differentiating children with Autistic Disorder from TD children at the end of the first year: Poor social interaction, lack of social smile, and lack of appropriate facial expressions. At the end of the second year eight IBSE items have consistently been identified as differentiating the two groups: Ignoring people, preferring aloneness, poor social interaction, lack of social smile, being too passive, lack of emotional expression, and increased distractibility. Zakian et al. (2000) also controlled for global developmental quotient (GDQ) and found no significant differences in behavioural signs on the IBSE for children with ASD and a GDQ of less than 50 compared to children with ASD and a normal GDQ (Zakian et al., 2000). This suggests that additional developmental delay does not impact on IBSE scores.

Wendland et al. (2010) found that the IBSE summary scores (based on 31 items; two could not be observed) were significantly higher for the ASD group during the 0-6 months and 6-12 months' time periods when compared to the TD group, indicating more abnormal behaviours. However, scores were no longer significantly different at age 12-18 months. This seems in contrast to the other studies using the IBSE which found an increase of number and / or severity of signs during the second year. However, Wendland's study sample was very small (n=6/group) and the results should therefore not be over-interpreted.

The study by (Adrien et al., 1992b) reported higher summary scores on the IBSE at all ages compared to the TD group, but did not test for statistical significance of these group differences. The ASD group scored particularly high on 11 items during the first year and 17 items during the second year, indicating an increase in problematic behaviours in the second year; statistical significance of the increase was also not tested. The videos in this study were assessed by coders with knowledge of the children's diagnostic status and thus the number of high scoring items may have been inflated.

Interactive Motor Evaluation Scales (IMES): The study by (Gattegno et al., 2005) (see also Adrien et al. (2005)) used the Interactive Motor Evaluation Scales (IMES), developed by the Tours research group to analyse videos of infants aged 4 to 6 months; the 36 item scale assesses interactive behaviours, joint attention and motor behaviours across four different interaction situations between infant and adults. Six behaviours were assessed for each of these situations, and five behaviours for when the child was alone. The behaviours were scored from 0 for absent or 1 to 3 according to their duration (1 to 5 or more seconds). A higher score indicates better functioning. This scale has not been validated yet and has not been used by other research teams. Children who later were diagnosed with autistic disorder (DSM-IV) had lower scores than typically developing children on all three domains of behaviour at age 4 to 6 months. The authors concluded that this indicated infants later

diagnosed with ASD show reduced interactive, joint attention, and motor behaviours. However, due to the small sample size, these differences were not tested for statistical significance.

Behavioural Functional Evaluation Inventory (BFE): The study by Malvy et al. (1994) used the Behavioural Functional Evaluation Inventory (BFE), which was also developed by the Tours research group. The BFI is a very detailed scale, comprising 104 items grouped into 13 functional domains; the scale's assesses impairments and abnormalities in social-communication behaviours, reactivity to sensory stimulation as well as motor behaviours. However, only four domains have been described in detail in the reviewed article and thus it remains what specific behaviours have been assessed. The items are scored from 0 for "abnormalities absent" or 1 to 4 according to the intensity of the behavioural abnormality. A higher score indicates more ore more severe behavioural abnormalities. A later version of the scale with 11 domains and 55 items has been validated (Adrien et al., 2001); however the version used in the study by Malvy and colleagues has not been validated. Malvy et al. (1994) found statistically significant differences between ASD and TD groups in the first and second year for scores across all sub-scales of the BFE. The same study also found a statistically significant increase in scores in most domains from the first to the second year for the ASD group. Only abnormal reactivity to sensory stimulation, passivity and low muscle tonus did not increase in severity during the second year.

Alarm Distress Baby Scale (ADBB): The study by Wendland et al. (2010) used the Alarm Distress Baby Scale (ADBB) in addition to the IBSE for analysis of home-videos. The ADBB was developed by Guedeney and Fermanian (2001) and assesses sustained infant social withdrawal, which is characterized by a decreased or lack of responsiveness to social stimuli. The scale assesses eight behaviours: facial expression, eye contact, activity, self-stimulation gestures, vocalisation to others, response to stimulation, capacity to engage with someone else, and capacity to demand someone else's attention. Each behaviour is scored on a 5 point Likert scale (0 to 4), with higher scores indicating more social withdrawal. A total score of 5 or higher is considered indicative of developmental risk. The scale was validated on a sample of 60 infants aged 2 to 8 months and showed sensitivity of 0.82 and specificity of 0.78 for detection of developmental risk (Guedeney & Fermanian, 2001). The scale has been used internationally in 13 controlled studies (Guedeney et al., 2013) of infant social withdrawal. Wendland et al.'s study found that the average sum score of children with ASD was above the ADBB cut-off score at 0-6 months, 6-12 months and 12-18 months, thus showing more social withdrawal behaviour than typically developing children (Wendland et al., 2010). The group difference in mean ADBB scores was large, but not statistically significant due to the small samples size (n=6/group) and the large standard deviation in the typically developing group. A statistically significant result would have required a sample size of at least 12 per group. Examining individual children's scores revealed that at 0-6 months only two children out of six children in the ASD group scored at or above the ADBB

cut-off, and at 0-12 months and 12-18 months only 3 out of 6 children in the ASD group scored at or above the clinical cut-off of the ADBB; this indicates considerable variability in social withdrawn behaviours among the ASD sample.

Infant Motor Maturity and Atypicality Coding Scale: Ozonoff et al. (2008b) developed the Infant Motor Maturity and Atypicality Coding Scale and assess videos of infants aged 9 to 13 months. The scale assesses developmental maturity of seven different postures or motor skills from least mature (=3) to most mature (=0), while noting the presence of any atypicalities in motor development. The scale has not been validated or used in any other research studies. Based on home-video observations, the authors found that children with regressive onset autistic disorder and children with other DD walked significantly later than TD children, that children with early onset autistic disorder and children with other DD had significantly delayed development regarding mature sitting than TD children. No differences between infants with ASD and typically developing infants were found in regard to atypical motor behaviours.

In summary, most evidence is available for the IBSE scale as it has been examined in four studies; for all other clinical scales the evidence was limited to one study each. Three of the five clinical rating scales (IBSE, BFE, IMMAC) used in home-video studies have found statistically significant differences either on individual items or on overall scale scores between children with autistic disorder and typically developing children. Two scales identified also an increase of number and severity of signs during the second year in children with autistic disorder (IBSE, BFE). The ADBB did not find any statistically significant differences in social withdrawn behaviour; however ADBB scores were significantly associated with IBSE scores which differentiated children with ASD from children with TD.

Table 2.3.1: Studies using clinical rating scales (N =7)

First author/ year	Country	Length of coded video	Age at recording (months)	# cases	# controls	Scale/ validated	Behaviours studied	Blind coders	Stat. analysis	QS
Adrien 1992	France (Tours)	10 – 80 minutes	0-12-24	AD: 11 DSM-III	TD: 3	IBSE / Yes	Social orienting, communication, motor behaviours, emotional response, perception/attention, adaptation to environment	No	No	5
Adrien 1993 & Adrien 1991	France (Tours)	10-80 minutes	0-12–24	AD: 12 DSM-III	TD: 12	IBSE / Yes	Social orienting, communication, motor behaviours, emotional response, perception/attention, adaptation to environment	Yes	Yes	11
Gattegno 2005 & Adrien 2005	France (Tours)	Not reported	4-6	AD: 9 DSM-IV	TD: 4	IMES / No	Interaction, joint attention, motor behaviours across 6 different situations	Yes	No	5
Malvy 1994	France (Tours)	Not reported	0-12-24	AD; 12 DSM-III	TD: 12	BFE / No	Attention, perception, association, intention, tonus, motricity, imitation, emotion, contact, communication, cognitions, instinct, regulation	Not stated	Yes	7
Ozonoff 2008	USA	118-138 min	9-13	23 AD-LO 25 AD-EO	24 TD; 25 DD	IMMACS / No	Maturity level of 5 motor behaviours: walk, crawl, sit, prone, supine. Movement abnormalities, lack of protective motor responses	Not stated	Yes	9
Wendland 2010	France (Paris)	Not reported	0-6-12-18	ASD: 6 DSM-IV	TD: 6	ADBB, IBSE / yes	ADBB: Social withdrawal behaviours IBSE: see Adrien et al 1992	1 blind, 1 informed	Yes	8
Zakian 2000	France (Tours)	Not reported	0-8-18-24	AD: 14 DSM-IV	TD: 10	IBSE / yes	Social orienting, communication, motor behaviours, emotional response, perception/attention, adaptation to environment	2 blind, 2 informed	Yes	12

AD: Autistic Disorder, ASD: Autism Spectrum Disorder; TD: Typical Development; IQ Ax: IQ assessment; QS: Quality Score

IMES: Interactive Motor Evaluation Scale; IBSE: Infant Behaviour Summarized Evaluation scale; ADBB: Alarm Distress de Bébé scale; BFE: Behavioural Functional Evaluation Inventory; IMMACS: Infant Motor Maturity and Atypicality Coding Scales

2.3.3 Evidence from studies using time sampled coding

All of time-sampled coding schemes used in the reviewed studies were developed for the purpose of the particular home-video study only and thus have not been validated. Three studies used coding grids that also allowed for rating of the quality of behaviours in addition to frequency and duration of occurrence (Bernard et al., 2005, 2006; Clifford & Dissanayake, 2008; Clifford et al., 2007). A broad range of behaviours has been captured by the different coding schemes including social communication and interaction behaviours, as well as motor behaviours and motor development. Statistically significant findings of these studies are summarised in the following paragraphs and presented in Table 2.3.4. The terminology used for target behaviours varied between studies; thus, it was attempted to match behaviours as much as possible based on the descriptions offered in these studies. For example, some studies categorised requesting gestures as communicative gesture, others as joint attention behaviour; in this review they have been summarised as joint attention behaviours.

Responding to name call by turning and looking to the person calling, has been consistently found in six out of seven studies to be reduced or lacking during the first year of life in infants later diagnosed with ASD or Autistic Disorder (AD) compared to typically developing infants (Baranek, 1999; Brisson et al., 2011a; Clifford & Dissanayake, 2008; Osterling, Dawson, & Munson, 2002; Osterling & Dawson, 1994; Werner, Dawson, Osterling, & Dinno, 2000) (not (Werner & Dawson, 2005). Furthermore, three of these studies with an additional control group of infants with developmental delay (DD) or intellectual disability (ID) have found that response to name call is more impaired in children with ASD than children with other DD (Baranek, 1999; Clifford & Dissanayake, 2008; Osterling et al., 2002). Thus reduced or lack of response to name call seems to be an ASD specific sign in the first year.

In the second year, the evidence for reduced response to name call in children with ASD remained consistent (4 out of 4 studies) (Clifford & Dissanayake, 2008; Clifford et al., 2007; Mars, Mauk, & Dowrick, 1998; Werner & Dawson, 2005). The majority of studies included children with a diagnosis of AD and children with a diagnosis of PDD-NOS in their samples, thus suggesting that impaired response to name call may be common across the autism spectrum. However, one study explicitly compared children with AD to those with PDD-NOS and found impaired response to name call only among infants with AD but not for infants with PDD-NOS (Mars et al., 1998). More research is needed to understand whether this is indeed a shared early sign for the whole autism spectrum.

Social smiling was not found to be reduced in infants with ASD in three out of fives studies during the first 12 months (Clifford & Dissanayake, 2008; Werner et al., 2000; Zappella et al., 2015); differences were found only at age 0 to 6 months in two of the studies

that examined children with AD only (Maestro et al., 2005a; Maestro et al., 2002). In the second year, three out of four studies reported no differences in social smiling between infants with ASD and TD infants (Clifford & Dissanayake, 2008; Clifford et al., 2007; Werner & Dawson, 2005); again the one study that reported a significant finding examined children with Autistic Disorder only (Maestro et al., 2006). Given the evidence, lack of or reduced social smiling appears to not be a reliable early sign of ASD, or may only be manifest in children with more severe ASD (i.e. AD).

Joint attention. Joint attention, i.e. the ability to share the same focus of attention with another person, was consistently found to be impaired in infants with ASD towards the end of the first year (5 out of 5 studies) (Clifford & Dissanayake, 2008; Osterling et al., 2002; Osterling & Dawson, 1994; Watson et al., 2013; Werner & Dawson, 2005). Two studies however, found that no impairments were present in infants with late or regressive onset of ASD during the first year (Osterling & Dawson, 1994; Werner & Dawson, 2005). Evidence for impaired joint attention behaviours in children with ASD was consistent in the second year (seven out of seven studies)(Clifford & Dissanayake, 2008; Clifford et al., 2007; Guidetti et al., 2004; Maestro et al., 2001; Mars et al., 1998; Watson et al., 2013; Werner & Dawson, 2005); three of these studies found that this differentiated children with ASD not only from children with TD but also from children with developmental delay (Clifford & Dissanayake, 2008; Clifford et al., 2007; Watson et al., 2013). Thus impaired joint attention appears to be a reliable early sign for ASD from about age 12 months onwards with the caveat that it may show later in children with late onset ASD.

Use of communicative gestures is closely related to joint attention behaviours; however communicative gestures including pointing, waving, nodding, shaking head have been often investigated separately to overall joint attention behaviours and thus are reported separately here. Us of communicative gestures has been consistently found (seven out of seven studies) to be impaired or reduced towards the end of the first year in infants later diagnosed with ASD (Colgan et al., 2006; Guidetti et al., 2004; Maestro et al., 2001; Osterling et al., 2002; Osterling & Dawson, 1994; Watson et al., 2013; Werner & Dawson, 2005). One study found that not the frequency of use but the repertoire of gestures was reduced in infants with ASD when compared to typically developing infants in the first year (Colgan et al., 2006). The study by Maestro et al. (2001) found that reduced pointing comprehension was already evident between age 0 to 6 months. One study found that reduced use of communicative gestures in the first year was only present in children with early onset ASD but not in those with late onset ASD (Werner & Dawson, 2005). During the second year evidence was limited to children with Autistic Disorder (AD) only but consistent for a reduced use of gestures in children with AD (five out of five studies) (Bernard et al., 2005, 2006; Clifford et al., 2007; Guidetti et al., 2004; Watson et al., 2013), with two of the studies also finding that this differentiated children with AD from those with DD (Clifford et al., 2007; Watson et al., 2013). One study found that only the frequency but not the repertoire

was reduced in children with AD during the second year (Guidetti et al., 2004). The combined use of gestures with gaze or vocalisation was also found to be impaired during the second year in children with ASD, however this was explored by a single study only (Bernard et al., 2005). It appears that deficits in gesture use appear to be a reliable early sign of ASD in the first two years; although it remains unknown whether these deficits still apply to the whole ASD spectrum in the second year as only children with AD have been examined.

Anticipation of others' actions or aim and attuning to these through gestures or postures, such as stretching up arms when about to be picked up or opening the mouth in anticipation of being fed, has been consistently found to be impaired in the first year of life in infants with ASD (4 out of 4 studies) (Baranek, 1999; Brisson et al., 2011b; Maestro et al., 2001; Maestro et al., 2002). In the second year this was only investigated by one study, which found that children with Autistic Disorder showed less anticipatory postures or gestures than children with TD as well as children with DD (Clifford et al., 2007).

Eye contact was only investigated by two TSC studies during the first 12 months: one study found reduced eye contact in infants with ASD when compared to typically developing as well developmentally delayed infants between age 0 to 12 months (Clifford & Dissanayake, 2008), while the other study did not find any differences between these groups at age 0 to 6 months (Zappella et al., 2015). During the second year evidence was also limited to two studies; these found that eye contact frequency and quality was reduced and that gaze aversion was present in children with ASD when compared to children with TD and DD (Clifford & Dissanayake, 2008; Clifford et al., 2007). It appears that impaired eye contact may be more indicative of ASD in the second year, and that this may include impairment in the eye contact quality rather than frequency only.

Looking at others. Five out of eight studies found that infants with ASD looked less at other people in the first year than typically developing infants. In two studies this difference was only apparent at age 0 to 6 months, but not at age 6 to 12 months (Maestro et al., 2005a; Maestro et al., 2002). In the second year, studies consistently found children with ASD were looking less at others than TD children (Maestro et al., 2006; Mars et al., 1998; Werner & Dawson, 2005). One study also found that if children with ASD looked at people they were less likely to look at their faces (Mars et al., 1998).

Vocalisations and language. Three studies consistently found that between ages 0 to 6 months infants with AD were directing vocalisations at others not at all or not as much as typically developing children (Brisson et al., 2011b; Maestro et al., 2005a; Maestro et al., 2002). The amount of any vocalisation overall however, was not found to be reduced in infants with ASD (Osterling & Dawson, 1994; Zappella et al., 2015); infants with early onset ASD had less complex babble than typically developing infants while infants with late onset ASD had more complex babble in the first year than TD infants (Werner & Dawson, 2005). In

the second year, TSC studies consistently found reduced or absent vocalisation overall (Bernard et al., 2005; Maestro et al., 2001; Maestro et al., 2006) and impaired or absent expressive language (Maestro et al., 2001; Mars et al., 1998; Werner & Dawson, 2005), including imitation of vocalisations in children with ASD when compared to TD children (Mars et al., 1998).

Affective expression was found to be reduced in children with ASD in the first year, but this was only investigated by one study (Baranek, 1999). In the second year, two studies found that children with AD showed reduced quality of affective expression compared to children with TD and DD (Clifford & Dissanayake, 2008; Clifford et al., 2007), but this was not confirmed in the study by Werner and Dawson (2005), which included children with a diagnosis of AD and PDD-NOS. Reduced affective expression may manifest more clearly in children with AD than in children from the less severe part of the spectrum (i.e. PDD-NOS).

Play behaviour. Play behaviours of infants with ASD were not found to be less developed or complex than that of children with TD or developmental delay (Baranek et al., 2005). In the second year, functional toy play was also not found to be impaired for children with ASD (Werner & Dawson, 2005), but engagement in social games was reduced (Clifford et al., 2007).

Physical contact. Only one study investigated physical contact seeking in the first year (Baranek, 1999), and only two studies investigated this in the second year (Bernard et al., 2006; Clifford et al., 2007). All three studies found that children with ASD did not seek out physical contact as much as TD children or avoided physical contact at all.

Deficits in other areas than communication and social interaction have been less well examined. These include the following behaviours:

Object exploration was found in two studies to be abnormal or reduced in infants with ASD during the first year of life (Baranek, 1999; Maestro et al., 2002);

Motor behaviours or development have been investigated in five TSC studies; of these three studies found differences between infants with ASD and TD infants for some motor behaviours including quality of general movements (i.e. arm, leg, torso movements) (Phagava et al., 2008; Zappella et al., 2015) and impaired movement symmetry (Esposito et al., 2009) during the first year. Motor development has been investigated by one study and found no delay for ASD infants in the first year, but in the second year (Lösche, 1990). Evidence for motor stereotypies being an early sign for ASD in the first year was supported by one study (Osterling et al., 2002), but another two studies found no differences in the presence of motor stereotypies between infants with ASD and TD infants (Werner & Dawson, 2005; Werner et al., 2000); none of the TSC home-videos studies have investigated presence of motor-stereotypies in the second year.

Overall, home-video studies using time-sampled coding grids have observed a broad range of behavioural differences in the first two years between children later diagnosed with ASD and typically developing children. The number of consistently reported behavioural signs was similar for both the first and the second year of life. In the first year reduced response to name call, reduced looking at others, reduced vocalisation to others, deficits in joint attention, impaired use of gestures, and deficits in anticipation of the aim or action of another person have been reported most consistently as signs for ASD. In the second year the signs most consistently reported included reduced response to name call, reduced looking at others, deficits in joint attention, reduced use of gestures, and reduced vocalisations overall as well as deficits in expressive language.

Table 2.3.3: Studies using time-sampled coding grids (N=24)

First author/ year	Country	Video length	Age at recording (months)	# cases	# controls	Behaviours studied	Coded for	'Blind' coders	Stat. analysis	QS
Baranek 1999 & 2005	USA	2 x 5 min	9-12	11 AD	10 DD 11 TD	1999: Affect expression, gaze, eye-contact, response to name, social touch response, motor stereotypies, object stereotypies, sensory modulation 2005: Object play behaviours	Frequency/ duration	Yes	Yes	11
Bernard 2005 & 2006	France	Not reported	12-17; 18-24; 24-29; 30-36	8 AD	8 TD	2005: Vocalisation, eye contact, use of gestures 2006: Use of gestures	Frequency and quality Frequency	Not stated	Yes	4
Brisson 2011a	France	Not reported	0-6	35 Autism	28 TD	Looking at others, smiling, vocalisation, response to name/ hello, briskness of response to social & non-social stimuli	Duration, reaction time	Not stated	Yes	5
Brisson 2011	France	Not reported	3-6	13 ASD	14 TD	Anticipatory mouth opening during feeding	Proportion of correct anticipation	Blind	Yes	9
Clifford 2007	Australia	5min / age	12-24	15 AD	15TD 15 Down Syndrome	Eye contact, contact seeking, response to name call, social smile, gestures, affect expression, peer interest, joint attention, play behaviour, social games	Frequency, quality	1 blind, 1 informed	Yes	11
Clifford 2008	Australia	10min/ age	0-12; 12-24	22 AD	19 DD 8 TD	Eye contact, affect expression, social smile, joint attention, requesting, social games	Frequency, quality	1 blind, 1 informed	Yes	9
Colgan 2006	USA	2 x 5 min	9-12	21 AD (DSM-III & IV)	14 TD	Use of gestures for social interaction, joint attention, behaviour regulation	Occurrence and function	Yes	Yes	11
Degenne 2009	France	Not reported	0-1; 4-5	5 ASD	4 TD	Gaze, head turning, vocalisation, motricity, facial expression	Frequency	Not stated	No	5

First author/ year	Country	Video length	Age at recording (months)	# cases	# controls	Behaviours studied	Coded for	'Blind' coders	Stat. analysis	QS
Esposito 2009	Italy	3 min	12-21 weeks	18 AD & PDD-NOS	18 TD 12 DD	Positional patterns for symmetry during lying	presence/ absence	Yes	Yes	11
Guidetti 2004	France	≥ 8min / age	12; 24; 36	6 AD?	6 TD	Communicative gesture, vocalisations and words; combination of both	Duration, frequency	Yes	Yes	7
Loesche 1990	Germany	18-69 min	4-12, 13-21, 22-30, 31-42	8 AD	8 TD	Stages 2-6 of Piaget's sensorimotor development; Spangler's 4 levels of action development (action-affect to goal oriented actions, and actions following a social script)	Frequency	Yes	Yes	5
Maestro 2001	Italy	1.5 hrs	0-6, 6-12, 12-18, 18-24	15 AD	15 TD	*Social behaviour:* looking at others, postural attunement, physical contact seeking, smiling, sharing enjoyment, vocalising *Intersubjectivity:* shared attention, pointing, imitating, anticipating other's action *Symbolic:* object exploration, symbolic play, communicative gestures, meaning vocalisation	Presence/ absence	Yes	Yes	8
Maestro 2002	Italy	10-62 min	0-6	15 AD	15 TD	Looking at people/objects; orienting towards people/objects; postural attunement, seeking contact, smiling at people/objects; attuning behaviour; vocalising to people/objects; anticipating other's action, object exploration	Presence/ absence	Yes	Yes	10
Maestro 2005	Italy	Not reported	0-6, 6-12	15 AD	13 TD	Looking at people/objects; orienting towards people/objects; smiling at people/objects; vocalising to people/objects	Frequency	Yes	Yes	11

First author/year	Country	Video length	Age at recording (months)	# cases	# controls	Behaviours studied	Coded for	'Blind' coders	Stat. analysis	QS
Maestro 2006	Italy	Not reported	0-6, 6-12, 12-18	15 AD-EO 15 AD-LO	15TD	Looking at people/objects; orienting towards people/objects; smiling at people/objects; vocalising to people/objects	Frequency	Yes	Yes	7
Mars 1998	USA	Not reported	12-30	25 AD & PDD-NOS	25 TD	Following verbal direction, looking at faces, showing of objects, alternating gaze, looking at people, combining pointing and gaze, words, verbal imitations	Frequency	Yes	Yes	11
Osterling 1994	USA	3-29 min	12	11 AD & PDD-NOS (DSM-III)	11 TD	Joint attention, response to name, seek contact, imitate, follow direction, speech, gestures, self-stimulatory behaviours	Presence/absence of behaviour	Yes	Yes	11
Osterling 2002	USA	Not reported	12	20 AD & PDD-NOS	14 ID 20 TD	Eye contact, joint attention, vocalisation, physical contact seeking, social games, imitation, response to name	Frequency, duration	Yes	Yes	10
Ozonoff 2011	USA	Not reported	6-24	52 AD	23 TD	Looking at people, smiling at people, vocalisation, joint attention(showing/requesting gestures)	Frequency	Not stated	Yes	8
Phagava 2008	Georgia	31sec – 3:50min	6-21 weeks	20 ASD?	20 TD	General movement patterns	Absence/abnormality	Yes	Yes	7
Watson 2013	USA	2 x 5 min	9-12;15-18	43 Autism	30 DD 36 TD	Use of gestures for social interaction, joint attention, behaviour regulation	Occurrence and function	Yes	Yes	10
Werner 2005	USA	Not reported	12; 24	26 AD & PDD-NOS	20 TD	Language, joint attention, gaze, orienting, repetitive behaviours, affect, toy play	Frequency, duration	Yes	Yes	10
Werner 2000	USA	2 – 38 min	8-10	15 AD & PDD-NOS (DSM-III)	15 TD	Language, joint attention, gaze, orienting, repetitive behaviours, affect, toy play	Frequency, duration	Yes	Yes	10
Zappella 2015	Austria/USA/Sweden	2 – 6 min	1-6	10 ASD	1TD 7 Tourette	General movements, concurrent motor repertoire, posture and tone, eye contact, responsive smiling, pre-speech vocalisations	Normal/abnormal	Not stated	Yes	5.5

2.3.4   Comparison of early signs detected by CRS and TCS home-video studies

Table 2.3.4 compares the results from studies using clinical rating scales (CRS) and results of home-video studies using a time-sampled coding approach (TSC). For ease of reporting the age periods were limited to 0-12 months and 12-24 months.

During the first year four behavioural signs have been most consistently reported across both study types; these included reduced response to name, reduced looking at others, and reduced or odd facial expression. Reduced use of gestures was very consistently reported by TSC studies for the first year, but was not investigated by any of the CRS studies. During the second year of life, five behavioural signs have been most consistently reported across both study types; these included: Lower quality of affect, reduced response to name call, reduced eye contact frequency, reduced looking at people, and reduced use of gestures. Reduced social smiling was consistently reported by CRS studies for the second year but not TSC studies.

Reduced vocalisation in the ASD group during the first year was more strongly supported by TSC studies than CRS studies; however CRS studies showed support for this difference during the second year; and both consistently found reduced imitation of vocalisation in the second year in children with ASD.

Joint attention was most consistently reported as being impaired in children with ASD by TSC studies in the first and second year; however, only two CRS study investigated this, of which one did not test for statistical significance (Adrien et al., 2005; Gattegno et al., 2005), and the other did not report on the join attention items but rather the overall clinical scale score (Wendland et al., 2010). Thus these were not included in the summary table.

There was some evidence from both study types for impairment in motor behaviours, however the different studies focussed on various different motor behaviours and thus no consistent evidence for one particular aspect of motor behaviours emerged. However, it appears that both study types consistently showed that stereotyped motor behaviours are not a distinguishing sign for children with ASD during the first two years of life.

Discriminant analyses were only conducted in six studies, all of them TSC studies; these reported rates for correct classification of between 78% and 94% for the first year, with ASD detection rates between 73% and 91%. These rates are considerably high, given that the only properly validated ASD screener for the first year, the parent-report based First Year Inventory, has sensitivity of only 0.44. In the second year, detection rates were reported by two TSC studies and these varied from 78% to 100%. Again these rates are high and similar to those reported for the SACS screener (see Chapter 1 for details).

Table 2.3.4: Evidence of early behavioural signs of ASD from home-video studies using clinical rating scales (CSR) and those using time-sampled observational coding grids (TSC)

Age in video:	0-12 months *		12-24 months	
Observational method:	Time-sampled coding	Clinical rating scales	Time-sampled coding	Clinical rating scales
	N of studies with significant results / total N of studies			
Affect behaviours:				
Reduced / odd facial expressions	1/1	2/3	-	2/3
Lower quality of affect	1/1	1/3	2/3	3/3
Interaction & Communication behaviours:				
Reduced / no response to name	7/8	1/1	4/4	1/1
Reduced eye contact frequency	1/2	0/2	2/2	1/2
Reduced eye contact quality	-	-	2/2	-
Reduced looking at people	6/8	1/1	3/3	1/1
Reduced / no social smiling	2/5	1/2	1/4	2/2
Prefers aloneness	-	1/3	-	3/3
Impaired joint attention	5/5	-	7/7	-
Anticipating others' actions	4/4	-	1/1	-
Reduced / avoidance of physical contact	1/1	-	1/2	-
Reduced/no vocalisation to others	3/3	1/3	-	3/3
Reduced vocalisation overall	1/3	-	3/3	-
Reduced/no imitation of vocalisations	-	1/3	1/1	2/3
Deficits in receptive language			1/1	
Lack of/ impaired expressive language	-	-	3/3	-
Reduced use of gestures	6/7	0/1	4/5	2/3
Motor behaviours:				
Decreased activity level	-	1/2	-	2/2
Hypotonia	-	1/2	-	1/2
Impaired general movements	2/2	0/1		-
Unusual postures	1/1	0/1	-	-
Repetitive motor actions / stereotypies	1/3	0/2	-	0/2
Delayed motor development	0/1	1/1	1/1	-
Other behavioural signs:				
Unusual object exploration	2/2	0/1		1/1
Impaired play behaviours	0/1	-	1/2	-
Higher distractibility	-	1/1	-	2/2
Social Withdrawal	-	0/1	-	0/1

*Zakian et al assessed signs at 0-8, 9-17 and 18-24 months, the findings for 9-17 months were omitted for this summary for ease of reporting. A dash '-' indicates this behaviour has not been examined.

2.4 Discussion and Conclusion

The aim of this systematic review was to synthesise the evidence for early behavioural signs of ASD from home-video studies, as well as compare the findings from home-video studies using clinical ratings scales (CSR) to those using a time-sampled coding approach (TSC).

Seven studies using clinical rating scales and 24 studies using time-sampled coding grids were identified. Home-video studies have investigated a wide range of behavioural signs, but only some have been investigated by more than one or two studies. Furthermore, many behaviours have been investigated by one study type only, or statistical analysis was only conducted by one type of studies (e.g. TCS studies for joint attention). For many behavioural signs, therefore, the existing evidence from home-video studies is insufficient to arrive at a confident conclusion about whether these are reliable markers for ASD.

A larger number of early ASD signs were reported for the second year of life than first year independent of the type of observation method used and this is in line with findings from previous reviews of home-videos studies as well as high-risk sibling studies (Jones et al., 2014; Mitchell et al., 2011; Palomo et al., 2006; Rogers, 2009; Saint-Georges et al., 2010). The number of signs consistently reported by both study types increased from three in the first year (response to name, looking at others, facial expression) to five in the second year (affect quality, response to name call, eye contact frequency, looking at others, use of gestures). One finding of this review is that reduced or lack of social smiling and eye contact appears not to be a distinguishing feature of children with ASD during the first two years of life. This is in line with the conclusions from earlier reviews of home-video studies (Palomo et al., 2006; Saint-Georges et al., 2010). However, there is some evidence from high-risk sibling studies that social smile may be reduced in children with ASD and differentiates them from children with DD as well (Jones et al., 2014). The current review also found no support for repetitive stereotyped motor behaviours being an early ASD sign, and this is concordant with findings from a systematic review of high-risks sibling studies (Jones et al., 2014).

For many of the early signs examined in the reviewed home-video studies it remains unclear whether these apply to the whole autism spectrum; most studies only included children with a diagnosis of Autistic Disorder (AD), the most severe disorder on the spectrum. However, one study has investigated differences between children with AD and children with PDD-NOS (Mars et al., 1998); their findings indicated that there may be a difference between these two groups. Thus it would be important for future studies to encompass the whole ASD spectrum, especially in light of the changed diagnostic criteria which no longer differentiate between AD, PDD-NOS and Asperger's Syndrome.

Challenges comparing both study types (TSC and CR) included the broad range of behaviours observed, and the differences in terminology used for referring to the same behaviours. For CSR studies an additional challenge was that some studies reported group difference on single item level while other only reported on difference in summary scores, making comparison of individual behaviours difficult. Another challenge was the different age brackets studies used for their observations. While some studies looked at one specific age or a fairly narrow age bracket (e.g. 12 months or 9-12 months), most studies made observations for larger age brackets, such as 0-12 months or 12-24 months; but not all studies reported on the mean or range of the age at time of recording; given the developmental milestones achieved during the second year of life, it would be expected that observations made closer to 12 months are different to those made closer to 24 months. This may explain to some extend the inconsistency in some of the findings.

Detection rates have not been widely reported I the existing home-video studies, especially not by studies using clinical ratings scales; CRS studies were also limited in their analysis by small sample sizes. Thus the usefulness of these clinical rating scales for the detection of ASD cannot be established from these home-video studies.

In conclusion, the reviewed home-video studies examine a wide range of early behavioural signs of ASD, and thus only few of these behaviours have been investigated across multiple studies, and even fewer across both video-analysis approaches (CRS and TSC). Nevertheless, there was some consistency in the findings of home-video studies using clinical ratings scales and those using time sampled coding grids. This suggests that clinical rating scales, and thus less detailed observations, may have potential to be clinically useful for early ASD screening. Overall, the review shows that a wide range of early deficits in social interaction and social communication behaviours can be observed in non-standardised settings; this means these signs may be most useful to focus on in ASD routing screening of infants and toddlers.

CHAPTER 3:   Methods

3.1   Feasibility of the study

The feasibility of the current study was tested prior to its implementation; an exploratory survey was distributed to parents within the researcher's social network and through the Autism Advisory and Support Service (AASS) to investigate availability of home video recordings for the relevant ages, as well as parents' willingness to share video- recordings of their children with researchers. Nine parents of children with ASD and 19 parents of typically developing children were surveyed online. Of these 93% had videos of their children when they were between 0 to 2 years old; 77% thought that it would be easy or very easy to locate these videos. Only parents of children with ASD were asked whether they were willing to provide their videos for research, all nine participants responded affirmative.

3.2   Sample size calculation

It was considered that a screening instrument for detecting signs of autism in very young children would only be clinically useful if it can show a considerably large enough difference between typically developing children and children with autism. Thus, a large effect size between the rating scores of cases and controls on the m-ADBB was desired. A large effect size (equivalent d = 1.0) will mean that those in the ASD group would score higher on the m-ADBB than those in the 'Control' group about 76% of the time when randomly selected (Fritz, Morris, & Richler, 2012).

In some instances, particularly in new research areas which do not have, for example, life-threatening implications, setting alpha at 0.1 can be more scientifically useful than the usual 0.05 (Lipsey, 1990). Lipsey (1990) argues that an alpha of 0.05 could too easily lead to a rejection of a 'true' finding (i.e. a false negative), and thus could incorrectly halt further research. Having a false positive at the early stage of the novel research is not so critical, however, as in this event continuing investigations will show whether the finding was indeed a false positive or not (by eventually applying the tighter alpha of 0.05). Based on this argumentation, and given the fact that the current study is only the second study to investigate social withdrawal as early indicator for ASD, alpha was set at 0.1 with a power of 80% and effect size at d=1.0 which resulted in a sample size of 10 participants per group for an independent sample t-test (Cohen, 1988). Because the type of distribution cannot be ascertained in small sample sizes, the non-

parametric equivalent of an independent sample t-test is required for analysis, which increases the sample size by 15% (Lehmann, 1998). Thus, a total of 11 to 12 participants per group were deemed necessary to have sufficient statistical power to detect a large effect size in m- ADBB score differences between the ASD and Control group. Based on the responses to the feasibility survey and the sample size calculations, as well as the assured support of an ASD organisation with over 600 parents accessing their services, it was deemed feasible to recruit the required numbers of 12 parents of children with ASD.

3.3   Statistical and clinical utility considerations

Given the anticipated small sample size of 11 to 12 participants per group, what results would indicate that the m-ADBB may be useful in the detection of children with ASD is essential?

First the statistical significance of the group difference of the m-ADBB scores between children with ASD and typically developing children will be tested. It was considered that children with ASD will have to score statistically significantly higher than children without ASD, and this difference must have a large effect size to be meaningful. As discussed in Chapter 1, usually, screening instruments are assessed on their sensitivity, specificity, positive and negative predictive value to determine their clinical usefulness. However, conducting a Receiver Operant Curve analysis was deemed infeasible for a small sample. Calculating these indicators for a small sample could therefore lead to misinterpretation of the results.

Secondly, the clinical significance of the findings also needs to be considered (Ogles, Lunnen, & Bonesteel, 2001). Given that this is a very new area of research, the investigators made an a-priori determination of what true and false detection rates would suggest that the m-ADBB might be clinically useful (see Table 3.3.1). The investigators agreed that if the m-ADBB did identify none of the children with ASD based on the cut-off of a total score of two or more, it would be deemed clinically not useful. Similarly, if the m-ADBB would flag all ASD and TD children as at risk for ASD, it would not be clinically useful either. Thus a clinically useful result was deemed to lie somewhere between these extremes.

Table 3.3.1: A-priori evaluation of clinical useful screening results in the current sample

Hypothetical screening results of the m-ADBB in a sample of N=11 per group		Researchers' judgement of clinical usefulness	
ASD	**TD**	**BC**	**SM**
0 to 2 detected	0 false positives	Not useful	Not useful
3 detected	0-2 false positives	Not useful	Possibly useful
4 or 5 detected	0-2 false positives	Possibly useful	Useful
6 detected	0-3 false positives	Useful	Useful
7 or more detected	0-4 false positives	Useful	Useful
7 or more detected	5-11 false positives	Not useful	Not useful

As can be seen, the criteria for what was considered clinically useful outcome differed slightly between both researchers, with the lower acceptable detection rate being three out of 11 and the upper limit for false positives being four out of 11. The acceptable number of false positives increased with the number of detected ASD cases.

If the results met theses determination, consideration will be made to surveying professionals in the field to ascertain their view on this topic. If however the data did not meet their determination (e.g., if none of the ASD children were detected on the m-ADBB) then no such survey would be required. This surveying of other health professionals is thus outside the scope of this thesis, but will be explored later if the results indicate that the investigators' determinations were met.

3.4 Research ethics

The study was approved by the Human Research Ethics Committee of the University of Sydney (Ref 2014/124) in April 2014, and modifications related to changed or new recruitment avenues were approved subsequently in 2014 and 2015. It is noteworthy that an amendment to approach private clinicians to advertise the study to parents of their clients was initially met with overwhelming resistance from the University Ethics Committee, despite this seeming to be a standard practice for recruitment of clinical study populations. The response was also surprising given that they had approved for ASD organisations to have their teaching or clinical staff promote the study to parents when they approved the original request. The Ethics Committee eventually approved recruitment through private clinicians after further negotiations. However, this led to a ten week delay for approaching private clinicians and their clients and for the subsequent recruitment through this channel.

Additional ethics approvals were obtained from Autism Spectrum NSW (ASPECT), KU Children's Services and Royal Far West for their staff to promote the study to parents of children with ASD (see Appendix F). All other organisations which were approached did not have a formalised approval process and agreed to advertising the study online or per flyer after submission of the research proposal and all relevant research documentation and Ethics approvals.

3.5   Participants

Parents were eligible to participate if their child was at least 2 years of age at the time of recruitment, and if they had video-recordings of their child at around age 12 months and / or 24 months, and if they spoke English. Children had to be at least two years old, as current evidence shows that this is the earliest age when a diagnosis of ASD can be reliably made (Ogles et al., 2001). Originally the lower age limit was set at three years but due to recruitment difficulties this was lowered to be more inclusive.

A total of 37 parents of children with ASD expressed interest in participating, of these 19 parents consented to participate, and 11 completed both online surveys and provided video recordings of their children. Most of the parents in the ASD group were recruited through autism organisations (n=7), the remainder were recruited through the researchers' social and professional networks (n=4). Twenty-one parents of typically developing children expressed interest in participating, of these 19 signed-up and 11 completed their study participation. Parents in the control group came from the researcher's own social network exclusively despite the study being advertised widely. The drop-out rate in both groups was 42% from consenting to participate to completion. See Table 3.5.1 for details. Reasons for non-completion are unknown as participants did not respond to follow-up emails. One of the parents who originally had signed-up as a control participant later informed the researcher that their child had subsequently been diagnosed with ASD (age of child at recruitment: 4 years 8 months). Thus this child was re-classified as an ASD participant, and the parent completed the additional ASD specific questions.

Table 3.5.1 Recruitment results

	Expressed interest N (%)	Consented N (%)	Dropped out N (%)	Completed N (%)
ASD group	37	19 (51)	8 (42)	11 (58)
TD group	21	19 (90)	8 (42)	11 (58)

ASD = Autism Spectrum Disorder, TD = Typical Development

The research study was promoted to parents through a variety of channels, including social media, online parenting forums, autism organisations, and private clinicians, as well as through the researchers' own social and professional networks. The overall reach is estimated to have been approximately 2000 parents of children with ASD, and approximately 500 parents from the general community, and more have been reached with untargeted advertisements on online parent forums, although this is impossible to quantify. Overall, the response rate was unexpectedly low; however sample sufficiently large to provide enough power for statistical analysis was recruited. See Appendix B for details of recruitment sources through which the study was promoted and the approximate reach for each source.

For 18 participants who consented but did not complete the study, information from the first survey was available. These did not differ significantly from participants by group membership (ASD vs. TD), parental age and sex, ethnicity or level of education; this suggests that drop-out of participants was non-selective on these demographic variables.

3.6 Procedure

Parents either expressed their interest to participate in the study by email or in person, where contact was made directly by the researcher (at AASS and KU Services). Once parents had expressed their interest, they were sent an email with information about what participation involved and the link to the first online survey. Parents were also able to directly register in the study online without prior contact with the researcher by following the link in the online advertisement distributed via different online parent forums or by email from supporting organisations. In this case parents were sent an information and welcome email after they had completed the first survey. Participants gave informed consent online after reading the information statement which was displayed at the start of the first online survey.

The first online survey screened for eligibility for participation as well as group membership by asking parents about the presence of a diagnosis of ASD for their child and to

rate their child's current social communication behaviour on a brief screener. Parents were also asked to provide a contact email address. Following the completion of the first online survey, participants were sent instructions per email for selecting and up-loading videos of their children to a secure online file sharing site (CloudStor). Participants then were asked to complete a second online survey, either the ASD or TD version according to their group membership. Once all videos were rated participants received individual feedback about the observations the researchers had made based on the video clips provided. Figure 3.6-1 visualises the overall study procedure.
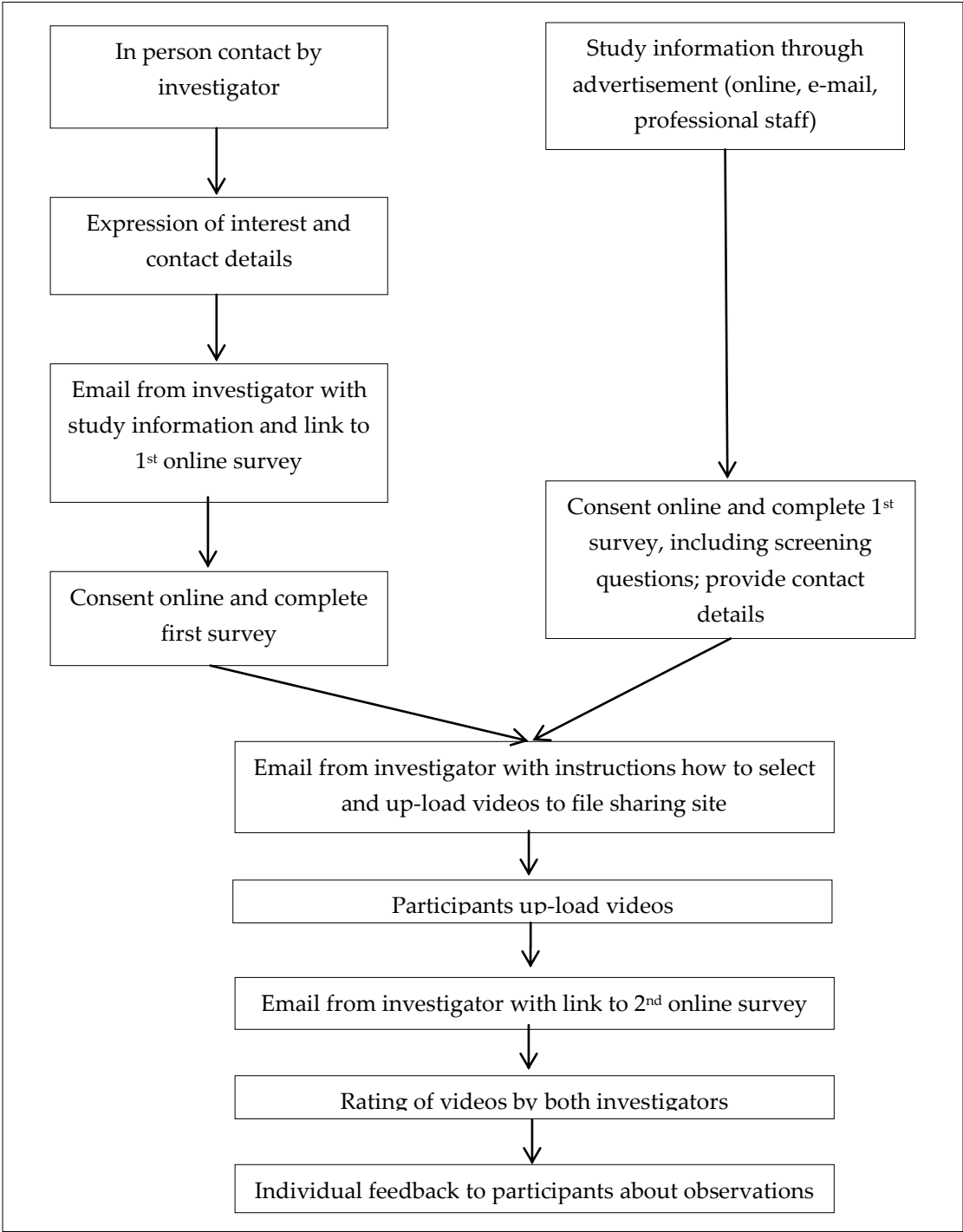
Figure 3.6.1: Study procedure

3.7 Measures

3.7.1 Online surveys

Two online surveys were designed and hosted on a commercial online survey website. The first survey included screening questions for eligibility, i.e. parents of a child aged 2 years or older, availability of video recordings from around age 1 and 2 years. The initial survey also screened for group membership by asking parents whether their child had been given a diagnosis of Autism Spectrum Disorder. Response options included Autistic Disorder, Asperger's, PDD-NOS and Autism Spectrum Disorder to ensure that all diagnostic labels from DSM-IV and DSM 5 were represented, as some children would have been diagnosed under the old criteria and some under the revised ones. As part of this survey, parents were also asked to complete a brief ASD screener, the brief version of the Social Responsiveness Scale (SRS-Brief) (Moul, Cauchi, Hawes, Brennan, & Dadds, 2014), with the aim to confirm group membership.

In the second online survey, parents were asked to provide details about the videos they had provided; e.g. a description of their child to ensure correct identification if more than one child was visible; date of recording; description of the setting or occasion of recording (e.g. family gathering, family routine, birthday), and whether other people visible in the video were well known to the child. This was deemed important to assess because children may behave more withdrawn around people they do not know well.

Questions about any problems with the child's hearing and vision from 0-2 years were also asked, to rule out that any observed difficulties in the child's social communication behaviour is caused by hearing and vision impairment. Parents were also asked about any concerns they may have about their child's development. Finally parents were asked whether they had experienced any mood difficulties in the two years following their child's birth, because parental depression can contribute to a child being socially withdrawn (see Guedeney et al. (2013)).

For parents of children with a diagnosis of ASD the second survey included additional questions about the diagnosis, including age of their child when first investigations began, who diagnosed the child and whether the health professionals involved in the diagnosis were part of a service specialised in childhood developmental disorders. It also asked how certain they felt that their child indeed had ASD, how they rated the impact of the ASD on their child's functioning, and whether they or anyone else had noticed early signs of ASD prior to their child's diagnosis and the type of signs they noticed. See Appendix D.3 for all questions included in the online survey.

3.7.2 Social Responsiveness Scale – Brief Version (SRS - Brief)

The SRS-Brief is a 16 item parents report scale (Moul et al., 2014); it has been developed from the original 65-item SRS (Constantino & Gruber, 2005). The original SRS was designed to assess autistic traits in children aged 4 to 18 years; however, the scale has been found to detect a broad range of behavioural problems in addition to ASD that also impair social communication (Towbin, Pradella, Gorrindo, Pine, & Leibenluft, 2005). To enhance the SRS' discriminant validity, Moul et al. (2014) conducted a factor analysis of SRS assessment results from different diagnostic groups, including ASD, Attention Deficit Hyperactivity Disorder, and Conduct Problems with and without callous-unemotional traits. A 16 item ASD factor was identified which best differentiated between these diagnostic groups. T-scores of 60 or more were found to be indicative of ASD. This 16 item ASD factor constitutes the SRS-Brief (see Appendix D.2). Due to its ease of administration and satisfactory validity, the SRS-Brief was used in this study to ascertain the group membership of participants, in particular to ascertain that control participants do not have impairments in their social communication abilities and are indeed typically developing.

3.7.3 The modified Alarm Distress Baby Scale (m-ADBB)

The modified Alarm Distress Baby Scale (m-ADBB; Matthey et al. (2013))) was used to rate the social interaction and communication behaviour of children as captured in the home-video footage provided by their parents. The m-ADBB is the modified Australian version of the ADBB, which is a French scale designed to screen for sustained social withdrawal behaviours in infants and toddlers (Guedeney & Fermanian, 2001). The original ADBB assesses eight behaviours with a 5-point scoring system for each. The modified Australian version assesses just five behaviours: facial expression, eye contact, vocalisation, activity and engaging in relationship with others; these behaviours spell the acronym FEVAR which can be easily memorised by clinicians. The items were reduced due to methodological issues with some of the items in the original scale, such as high inter-correlation or difficulties in achieving satisfactory inter-rater reliability. Each behaviour on the m-ADBB is rated as 'satisfactory', a 'possible problem' or 'definite problem', with clear descriptors for each scoring option. Vocalisation is rated as either satisfactory or possible problem only. The presence of at least two possible problems or one definite problem indicates social withdrawal (see Appendix D.1). The scale has been validated for use by nurses at early childhood clinics or other health professionals, and suitable for infants/toddlers 3-24 months old and has good concordance with the full ADBB scale (Matthey et al., 2013). The original ADBB scale has been used in 13 controlled studies across nine countries examining infant social withdrawal (Guedeney et al., 2013). It has also been used in one home-video study of children with ASD (Wendland et al., 2010). Less research has been conducted with the m-ADBB (Hartley et al., 2010; Matthey et al., 2013; Matthey et al., 2005).

3.8   Video material

The aim was to collect video clips of at least three minutes duration for each age and participant. Participants were instructed to provide video clips that showed someone interacting with, or being present with their child for most of the time; that is, the child was not video recorded just playing by himself/herself with no-one else being present but rather with someone present trying to engage him/her on occasions.

All videos were edited by a third person not involved in rating the video to avoid raters to be biased. Videos were edited to remove any footage that did not show the child and to mute comments by parents that would give away the group membership of the child. Where parents provided more than one video clip for the same age these were merged to provide sufficient viewing material. The video editor also assigned new IDs to the videos to avoid any incidental recognition of IDs as belonging to the case or control group by the researcher who conducted the recruitment. Furthermore, the videos for both time points (12 and 24 months) for the same participant were given non-matching IDs in order to avoid knowing the participant's group membership through accidental recognition of IDs after rating the 12 months videos. In addition, all 12 months videos were watched and rated before the 24 months videos as it was deemed more likely for behavioural signs to be observable at 24 months than 12 months. Thus, if the researchers would recognise a child in the 24 months videos from the 12 months videos, there would be less chance of "halo effect" from "problem" ratings at 12 months.

The author of this thesis acknowledges that she was not blind to five children's group membership prior to watching their videos, because some of them were recruited from her own social network and thus they were known to her. The second researcher (SM) however was blind to the group membership of all children. Whether ratings were less or more likely to be concordant between both raters (BC and SM) when these children were known to researcher BC was examined and results are presented in the next chapter.

3.9   Video rating procedure

Version 32 of the m-ADBB was used for rating children's behaviour captured in the video-recordings provided by their parents (see Appendix D.1). Training in the use of the m- ADBB was provided by the primary research supervisor and author of the scale. Reliability accreditation was achieved prior to rating the research video footage. Reliability accreditation was given once 80% concordance in ratings of six reliability video clips was achieved.

Videos were rated in blocks of four by each researcher separately. Both researchers' ratings were then compared to ensure ongoing reliability and prevent rater drift. Where there

were discrepancies, both researchers met, viewed the video together, and discussed the discrepancies. As part of these discussions the definition for brief eye contact was altered, and a more lenient definition was adopted (less than one second instead of one second duration); the m-ADBB scoring instruction for eye contact was subsequently changed on the rating sheet; this new version (v 32) was then used for all subsequent video ratings and previous videos were re-rated for eye contact if they were previously rated as less than satisfactory on this item. Overall, when there were differences in ratings these were resolved easily and were mostly due to one rater having missed a particular behaviour; e.g. very brief facial expression or eye contact, rather than these behaviours being interpreted differently.

It was considered possible that both researchers may pick up on other clinically relevant signs than those rated by the m-ADBB that may convince them of a child's group membership After all videos were rated, both researchers therefore used their clinical judgement in addition to the observed five m- ADBB behaviours to categorise each child as either having ASD or being typically developing. Only after completion of all ratings and clinical judgements were group-memberships revealed.

3.10 Data Analysis

All data from the surveys were downloaded into SPSS and merged into one database. Video ratings for 12 and 24 month videos were entered manually into the merged database. The overall m-ADBB ratings were then converted into scores. For each item a 'Satisfactory' rating was assigned a score of 0, a ratings as 'Possible problem' was assigned a score of 1, and a rating of 'Definite problem' was assigned a score of 2, creating a possible score range from 0 to 9. Where an item was considered borderline between 'Satisfactory' and 'Possible problem', a decision was made to code that item as 'Possible Problem', as this item was clearly considered a not being perfectly satisfactory, and to avoid fraction of scores which would make reliability more difficult to achieve. Thus the cut-off of at least one definite problem or two possible problems on the m-ADBB equates to a total score of 2 or more.

Inter-rater reliability (IRR) was examined for the m-ADBB classifications (below cut-off or equal or above cut-off for social withdrawal), the m-ADBB sum scores, the individual m-ADBB items and the clinical judgements of group membership. To establish IRR percentage agreement as well as intra-class correlations (ICC) were calculated (Hallgren, 2012). Descriptive statistics were computed for all sample characteristics.

To test for significant group differences in the m-ADBB scores and other interval and ordinal scale variables Mann-Whitney U-tests and for group comparison of categorical data chi-

square tests were conducted. Significance levels were set to $p \leq 0.1$ for all tests, except for group comparisons for individual m-ADBB items, which were Bonferroni adjusted for multiple tests (n=10) , and thus the significance level required was $p \leq 0.01$.

Effect sizes were calculated for all statistical tests: for Mann-Whitney U-tests Cohen's 'r' was calculated and for chi-square tests the effect size ϕ was calculated (Fritz et al., 2012). An effect size is classified as large if r or $\phi \geq 0.5$, medium if r or $\phi \geq 0.3$, and a small if r or $\phi \geq 0.1$ (see (Fritz et al., 2012)). The formula for these effect sizes are as follows (Fritz et al., 2012):

$$r = \frac{z}{\sqrt{N}}$$

$$\phi = \sqrt{\frac{\chi^2}{N}},$$

All data analyses were undertaken with SPSS version 22 (IBM Corp., Released 2013), with exception of effect size calculations which were undertaking in Microsoft Excel.

CHAPTER 4: Results

This chapter presents the results of the study. Sample characteristics, diagnostic information about the ASD cases and a description of the video material are presented first, followed by the analysis of the inter-rater reliability, statistical analysis of group differences in the m-ADBB scores and an examination of the detection rate in the sample.

4.1 Sample Characteristics

Group membership was ascertained with a brief parent report questionnaire assessing the children's social responsiveness at the time of recruitment (SRS-Brief; Moul et al. (2014)). All children in the ASD group scored within the clinical range for ASD and all children in the TD group scored below the clinical range for ASD on this measure.

The sample consisted mostly of highly educated parents of Caucasian ethnicity. Parents in the TD group had a statistically significantly higher level of education than parents in the ASD groups (X^2=8.25, p=0.041, ϕ=0.61). The gender distribution of participants' children was similar in both groups (X^2=0.79; p=0.375; ϕ = 0.19). Children in the ASD group were on average older than children in the TD group at time of recruitment, but the difference was not statistically significant (Z=-1.67, p=0.094, effect size r=-0.37).

Children in the ASD group had statistically significantly lower parental rating of their overall health in the first two years of life than children in the TD group (X^2=7.27, p=0.026; effect size Φ= 0.57). Significantly more parents in the ASD group reported developmental concerns in addition to their child's ASD compared to parents in the TD group (X^2= 10.78, p=0.002, ϕ=-0.7). In the ASD group developmental concerns included concerns about language development (n=4) and motor development (n=2), auditory processing difficulties (n=1), aggression (n=1), feeding (n=1) and sleeping difficulties (n=1). In the TD group two parents reported concerns about their child's language development. See Table 4.1.1 in for more details on participant characteristics.

Table 4.1.1: Participant characteristics by group membership

	ASD (N =11)			TD (N=11)			Test statistic	Effect size[†]
	N	%	M (SD)	N	%	M (SD)		
Parent characteristics:								
Age at recruitment (years)			38.2 (6.2)			35.9 (2.6)	Z^b =-1.80	r= -0.36
Sex of informant:								
Female	9	82		11	100		X^2 =2.2	Φ= 0.19
Male	2	18		0	0			
Ethnicity:								
Caucasian	10	91		9	82		X^2 =0.39	Φ= 0.13
Other	1	9		2	18			
Highest level of education:								
University	5	46		11	100			
TAFE	3	27		0	0		X^2 =8.25*	Φ= 0.61
Year 12 or HSC	2	18		0	0			
Year 11 or below	1	9		0	0			
Mood difficulties ≥ 2 weeks in first 2 years after birth	6	55		3	27		X^2 =1.69	Φ= 0.28
Child characteristics:								
Age at recruitment (years)			5.9 (3.5)			3.5 (1.9)	Z^b =-1.67	r= -0.36
Sex:								
Female	3	27		5	46		X^2 =0.79	Φ=0.19
Male	8	73		6	55			
SRS-Brief screening results:								
Score > ASD cut-off [a]	11	100		0	0		X^2 =22.00**	Φ= 1.00
T-score			99.6 (17.7)			40.0 (3.4)	Z^b =-3.98**	r= -0.85
Hearing or vision problems:								
At age 12 months	3	27		3	27		X^2 =0	Φ= 0
At age 24 months	3	27		1	9		X^2 =1.22	Φ= 0.24
Overall health in first 2 years:								
Very good	3	27		8	73			
Good	3	27		3	27		X^2 =7.27*	Φ= 0.57
Fair	5	46		0	0			
Poor	0	0		0	0			
Developmental concerns [c]	10	91		2	18		X^2 =10.78**	Φ= 0.70

ASD= Autism Spectrum, TD= Typical Development; [a] T-score ≥ 60 indicates ASD; [b] Mann-Whitney U test for independent samples; [c] other than ASD; *p<0.05; **p<0.01; [†] Effect sizes: ≥0.1 small, ≥ 0.3 medium, ≥ 0.5 large

4.2 Parent report of ASD diagnosis and early signs

All children were reportedly diagnosed by health professionals: five by a paediatrician, one by a paediatric psychiatrist, and the other five through a specialist Child Development Unit. The earliest diagnosis was given at 23 months of age (ASD4) and the latest was 10 years (ASD10), half of the children were diagnosed by age 45 months; the average age of diagnosis was 51.4 months (SD 28.5 months), which is similar to the average age of diagnosis reported in the literature (Chakrabarti & Fombonne, 2001; Fombonne, 2003). Two parents reported a diagnosis of PDD-NOS and one a diagnosis of Asperger's Syndrome for their children, all others reported that their children have been diagnosed with ASD, and four of these reported an additional diagnosis of global developmental delay. See Table 4.2.1 for details.

All but one parent reported having noticed early signs, between 4 months to 40 months; half had noticed signs by age 15 months, with an average age of 16.3 months (SD= 11.6); this is six months later than what previous research has reported as an average age when parents notice first signs (Young et al., 2003). One parent reported that childcare staff first noticed signs of developmental difficulties around age 18 months.

Signs that parents had noticed included delayed language development (n=6), child not pointing when requesting something (n=4), not responding to name call (n=4), lack of social smile (n=3), lack of eye contact (n=3), meltdowns (n=3), hypersensitivity to certain sensory inputs (n=3), aggression (n=2), self-injurious behaviours (n=1), not seeking out affection (n=1), and sleeping difficulties (n=1). All participating parents reported noticing more than one sign. Furthermore six parents reported that their child had lost skills he/she had previously developed, such as language skills (n=5) and making eye contact (n=3). The age when these six parents noticed the loss of skill was between age 12 months and 24 months (mean=16.9; SD=4.5).

Comparing the age of diagnosis with the age when parents had noticed first signs shows that the participants' children had received a diagnosis of ASD between 5 months and up to 8.5 years after their parents had noticed first signs (median= 24 months; mean=35.0, SD= 29.4).

Nine of the 11 parents of children with ASD stated that they would have liked to know that their child had ASD earlier than when the official diagnosis was given. The impact of ASD on their child's current functioning was rated by most parents as moderate (n=8), two rated it as mild and one as severe. Detailed characteristics of each of the children in the ASD group are presented in Table 4.2.1 together with their m-ADBB ratings.

Table 4.2.1: Characteristics of ASD sample according to parent report and m-ADBB ratings

ID	Sex	Diagnosis	Age at diagnosis (months)	Loss of skills	Age of first signs (months)	Level of impairment	SRS-Brief T-Score[#]	12 month	24 month
			Parent report					**m-ADBB ratings**	
ASD1	M	ASD & Global DD	56	No	4	Mild	98	Problems	Problems
ASD2	M	ASD	30	Yes	20	Moderate	102	Problems	No problems
ASD3	M	ASD & Global DD	30	No	6	Moderate	98	n/a	No problems
ASD4	M	PDD-NOS	23	Yes	4	Severe	116	No problems	n/a
ASD5	M	PDD-NOS	42	Yes	15	Moderate	65	Problems	No problems
ASD6	M	ASD	45	No	24	Moderate	94	Problems	Problems
ASD7	M	ASD & Global DD	45	Yes	40	Moderate	83	No problems	Problems
ASD8	M	ASD	28	Yes	13	Moderate	113	No problems	No problems
ASD9	F	Asperger's	69	No	30	Mild	89	No problems	No problems
ASD10	F	ASD	120	Yes	18	Moderate	107	No problems	No problems
ASD11	F	ASD & Global DD	77	No	6	Moderate	131	Problems	Problems

ASD = Autism Spectrum Disorder (DSM 5); DD = Developmental Delay; PDD-NOS= Pervasive Developmental Disorder Not Otherwise Specified (DSM-IV); [#] T-score ≥ 60 indicates ASD; SRS-Brief = Social Responsibility Scale-Brief version; m-ADBB= modified Alarm Distress Baby scale; n/a = not applicable because no video was available for this age

4.3 Video material

Video recordings were available for 22 participants; 19 participants had videos for both time points (12 and 24 months); one participant had provided video recordings for 12 months only, and two participants had video clips for 24 months only. One 24 months video was excluded as it was too brief (19 seconds) to allow for valid observations (ASD4). As result, 10 videos were available for each group and each time point and 40 video recordings overall. All videos had sound. Details of all video recordings are presented in Table 4.3.1.

The average lengths of video recordings at 12 months was 6:09 minutes (SD=5:17; range: 0:46 – 16:02 minutes) for the ASD group and 2:52 minutes (SD=3:21; range: 0:36 – 11:28 minutes) for the TD group. At 24 months the average lengths of videos was 5:46 minutes (SD=7:06; range: 0:37 – 22:31 minutes) for the ASD group and 2:39 minutes (SD=2:28, range: 0:49 – 9:13 minutes) for the TD group. Mann-Whitney U tests showed that the video length in both groups was not statistically different at 12 months (Z=-1.74, p=-0.082; ɸ=-0.39) or at 24 months (Z=-1.21; p=0.23; ɸ=-0.27) with small to moderate effect sizes. Overall, 12 months videos were on average longer than 24 months videos in the whole sample but this difference was not statistically significant and the effect sixe was moderate (Wilcoxon Signed Ranks: Z=-1.85; p= 0.064; ɸ= -0.41).

In the 12 months videos, children in the ASD group were on average 11.9 months old (SD=1.3) and children in the TD group were 12.5 months old (SD=1.7). Mann-Whitney U tests showed this difference was not statistically significant and the effect size was small (Z=-0.27; p=0.786; effect size r=-0.06). In the videos provided for the 24 months age, children in the ASD group were on average 23.9 months old (SD=2.7) and children in the TD group were on average months 24.1 months old (SD=2.0); this difference was not statistically significant either (Z=-0.611; p=0.541; effect size r=-0.14).

The videos showed the children in different settings, but mostly within the family home interacting with their parents, siblings or grandparents either in play or as part of family routines. Some of the videos showed the child at family gatherings such as child's birthday or at Christmas. Thus, the situations in which children were filmed appeared to be similar for both groups. The videos were also assessed for the number of adults and children who were present in the video in addition to the index child. As can be seen in Table 4.3.1, in both groups there were three participants with 12 months videos and four participants with 24 months videos, in which the child and the camera person were by themselves; in all other 12 and 24 months videos at least one other person was present. Thus, the videos were comparable between both groups in showing opportunities for interaction for the index child. All parents in the ASD and TD group reported that all people visible in the videos were reasonably well known to the child;

with exception of the 12 months video for TD3, were only some people were reasonably well known to the child.

To determine any selection bias parents might have had when choosing the video recordings they were asked to briefly state the reason for their selection. The most frequent reason reported was that videos selected were the only ones meeting the selection criteria (N=10); i.e. being recorded around the required ages of 12 months and 24 months, showing the child interacting with someone and being of sufficient duration. Two parents of children with ASD chose videos because they thought these recordings would show early behavioural signs of ASD. Overall, the video recordings could be considered to be comparable between both groups in regard to child age, video length and content.

Details of the length, content of the video recordings and number of people present in each video are presented in Table 4.3.1. Where there were multiple video clips available, information is presented sequentially for each clip. For example, for ASD2 the '12 month' video showed two different scenes– one when the index child was 12 months old, and the second when he was 15 months old. The first was of the index child at child care, with two adults and one other child in the clip, and the second with just one adult present (the parent with the camera).

Table 4.3.1: Characteristics of participants' video recordings

	12 month video				24 month video			
ID	Age (months)	Duration (mm:ss)	Type of scenes	Nº of people [†]	Age (months)	Duration (mm:ss)	Type of scenes	Nº of people [†]
ASD1	10	02:30	Playing at home	2A	24.5	01:36	Family routine	2A
ASD2	12 / 15	02:14	At child care/ playing at home	2A,1C/ 1A	18/20	02:50	Playing at home	1A
ASD3			*No video available*		24	22:31	Playing at home/ Family routine	1A/2A
ASD4	10 /11	03:16	Playing at home/ at restaurant	1A	28	0:19[‡]	Playing at home	1A
ASD5	12	08:02	1st birthday party/ family routine	5A/1A	23	05:40	Playing at home	2A
ASD6	12 / 13	16:02	1st birthday party/ family routine	2A,9C/ 1A,2C	22/24	14:30	Playing at home/ family routine/ birthday	1A,2C/ 1C,1A/ 2A,3C
ASD7	13	09:57	Playing at home	2A,1C	23/26	2:39	With guests at home/ family routine	4A,1C/ 4A
ASD8	13	00:46	Family routine/ playing at home	1A	22	00:37	Family routine	1A
ASD9	12	04:20	Family routine/ playing at home/playground	2A/1A/ 1A,1C	30	01:00	Playground	~10A, ~8C
ASD10	12	12:55	Family routine/ family Christmas	1A/6A, 1C	24	03:36	Family routine/ playing at home	2A/1A
ASD11	10	01:32	Family routine/ playing at home	1A	24	02:47	Playing at home	1A
TD1	11	02:16	Family routine	2A	22	03:53	Playing at home	2A
TD2	12	04:50	Family gathering/ playing at home	4A/1A	23/24	02:36	Playing at home/ with guests at home	1A/~3A

	12 month video				24 month video			
ID	Age (months)	Duration (mm:ss)	Type of scenes	Nº of people [†]	Age (months)	Duration (mm:ss)	Type of scenes	Nº of people [†]
TD3	12	03:28	1st birthday party/ playing at home	4A/1A	*No video available*			
TD4	12	00: 45	1st birthday party/ playing at home, family routine	2A,1C	22/25/27	02:20	Playing at home / playing at home / in the park	1A/ 1A,1C/ 1A,1C
TD5	11	03:07	Playing at home	1A	23/27	01:13	Family routine / at child care	1A
TD6		*No video available*			23-32	02:13	At playground/ 2 scenes playing at home / child care	1A/2A/ 1A,1C/ 2A,2C
TD7	11 / 14	11:28	Playing at home	1A,1C	26	09:13	Playing at home	2A,1C
TD8	15	00:40	Playing at home	2A	19/24	01:08	On the street / playing at home	2A/1A
TD9	12	00:54	Playing at home	1A	24	00:49	Playing at home	1A
TD10	14 / 18	00: 45	Playing at home	1A	23/25	01:40	Family routine	1A
TD11	11	00:36	Playing at home	2A	24	01:34	Playing at home	3A

ASD = Autism Spectrum Disorder, TD= Typical Development; [†] number of people per scene other than index child, but including the person behind camera, estimations indicated by '~'; A =Adult; C = Child; [‡]this video was not rated due to its brevity

70

4.4 Inter-rater reliability for m-ADBB ratings

To assess the degree that ratings were consistent between both investigators, for both the 12 months and 24 month videos, two-way mixed, accuracy, average-measure intra-class correlations (ICC) (McGraw & Wong, 1996) where calculated. Percentage of agreement was also calculated for better understanding of the inter-rater reliability, as ICC values can be affected by the degree of variance of the scores, and some items had low variance in their scores. All calculations were based on ratings before consensus was reached between investigators on any discrepancies between ratings.

The concordance for m-ADBB categorical ratings (problems vs. no problems) is illustrated in Table 4.4.1. It shows that both investigators gave more "no problems" ratings to children than they gave ratings indicating problems; the agreement for "Problems" ratings was lower than for "No Problems" ratings.

Table 4.4.1: Comparison of m-ADBB ratings[1] of both investigators

	Investigator 2 (BC)*		
	No problems	Problems	Total
Investigator 1 (SM)	12 month videos		
No problems	**12**	3	15
Problems	2	**3**	5
Total	14	6	20
% Agreement			15/20 = 75%
	24 month videos		
No problems	**13**	2	15
Problems	1	**4**	5
Total	14	6	20
% Agreement			17/20 = 85%

* Group membership of five children was known to Investigator 2; see section titled 'Knowledge of group membership'; rater agreement in **bold**

[1] Refers to 2-group classification of 'problems' and 'no problems'

Table 4.4.2 shows the ICCs as well as percentage agreement for the m-ADBB two-group classifications, the total sum scores and for individual m-ADBB item scores for both ages. ICC values between 0.40 and 0.59 are considered fair; values between 0.60 and 0.74 are considered

good; and values between 0.75 and 1.0 indicate excellent inter-rater reliability (Hallgren, 2012).

IRR for 12 months videos. For the 12 month videos, percentage agreement ranged from 60% to 90%. The agreement for overall m-ADBB classification was 75%, while the agreement on m-ADBB sums scores was lower with 60%. When examining agreement for individual items, the highest agreement was reached for 'Activity' (90%) and the lowest for 'Eye contact' (60%). The ICCs for 12 months indicate fair agreement for m-ADBB classifications and m-ADBB sum scores. Good agreement was achieved for the items 'Facial expression', 'Eye contact' and 'Vocalisations'. The ICC for Activity could not be calculated because almost all children scored 'satisfactory' on this item, thus the score variance for this item was close to zero.

IRR for 24 months videos. The IRR for 24 month video ratings was overall higher than for 12 months video ratings. The ICC for overall m-ADBB ratings indicated good agreement, and excellent agreement for m-ADBB sum scores. The items 'Facial expression' and 'Eye contact' yielded the highest ICCs at 24 months. For 'Activity' the ICC could not be calculated as all subjects were assigned the same score ("satisfactory") by both investigators, thus the variance scores for this item was zero. This indicates that the item 'Activity' may not be useful in differentiating children in the ASD group from those in the TD group at this age.

Table 4.4.2: Inter-rater reliability for m-ADBB ratings

	12 month			24 month		
	% agreement	ICC		% agreement	ICC	
Overall m-ADBB classification[1]	75	0.56	fair	80	0.66	good
m-ADBB sum score	55	0.45	fair	60	0.86	excellent
m-ADBB items:						
Facial expression	70	0.66	good	90	0.91	excellent
Eye contact	65	0.63	good	85	0.82	excellent
Vocalisation	70	0.66	good	90	0.64	good
Activity	90	n/a	-	100	n/a	
Relationship	75	0.46	fair	75	0.78	excellent

[1] Refers to 2-group classification of 'problems' and 'no problems'; n/a= this item score had zero or very low variance, ICC cannot be calculated; *not valid due to limited variance

Knowledge of group membership. Because participants were also recruited through Investigator 2's own social network, the group membership of five of the children was known to her. To assess whether Investigator 2's knowledge of group membership in five cases had impact on inter-rater reliability, percentage agreement was calculated separately for ratings of these five children. Agreement between both investigators was 80% at both 12 months and 24 months for the overall m-ADBB rating and m-ADBB sum scores. This indicates a somewhat higher IRR at 12 months for cases that were known to one of the investigators when compared to the overall sample, and a similar IRR at 24 months with the exception of m-ADBB sum scores. ICC was not calculated due to the very small sample size. In conclusion, one Investigator 2's knowledge of the five children's group membership may have slightly inflated IRR at 12 months, but not at 24 months. Thus, Investigator 1, who was uniformed of the group membership of all children, gave very similar ratings for these five cases as Investigator 2; and these ratings were at the same level of agreement as the ratings for the children whose group member status was unknown to both investigators. Thus, while Investigator 2's knowledge of group membership in five cases is a limitation of the study, it does not appear to have resulted in biasing the ratings.

Consensus for discrepant ratings. Where ratings differed between investigators, a consensus rating was reached after both investigators watched the video together and discussed discrepancies. Consensus was reached without difficulty in all cases; discrepancies were mostly due to one investigator having missed behaviours of interest when these were displayed only very briefly by the child (e.g. a very brief facial expression). The consensus ratings were used in all further analyses.

4.5   Group differences in m-ADBB sum scores

Total m-ADBB scores. It was hypothesised that children with ASD will score statistically significantly higher than TD children on the m-ADBB at both ages. Therefore, the null-hypothesis was that there will be no statistically significant difference in the m-ASBB scores between the ASD and TD groups at both ages. To test this hypothesis, a Mann-Whitney U-test for independent samples was conducted. Results including effect sizes are presented in Table 4.5.1. See Chapter 3 for details on effect size calculation. An effect size 'r' of 0.1 or higher is considered small, 0.3 or higher is considered medium, and 0.5 or higher is considered large. For negative 'r' values the absolute value without regard to the sign is interpreted.

Table 4.5.1: Mann-Whitney U test for group comparison of m-ADBB sum scores

Group	N	Mean	SD	Median	Mean Rank	Sum of Ranks	Z	p-value	Effect size r
~12 month									
ASD	10	1.4	1.3	1.5	13.45	134.50	- 2.54	0.023	- 0.57
TD	10	0.2	0.6	0.0	7.55	75.50			
~24 month									
ASD	10	1.7	1.4	1.0	13.50	139.50	- 2.40	0.023	- 0.54
TD	10	0.7	1.6	0.0	7.50	70.50			

Z = Mann-Whitney U-Test statistic; SD = Standard Deviation

The mean rank for the ASD group was statistically significantly higher than those for the TD group in the 12 month videos and the 24 month videos, and this difference had a large effect sizes ($r \geq |0.5|$). Thus, the null-hypothesis, that the m-ADBB scores for children with ASD are not different from those of TD children, can be rejected, and the results lend support to the research hypothesis, that children with ASD score statistically significantly higher on the m-ADBB than TD children in their 12 month and 24 month videos.

Individual m-ADBB items. After adjusting for multiple testing (Bonferroni method) none of the group differences for the individual m-ADBB item scores were statistically significant at age 12 month and age 24 month. However, at 24 month the group difference for the item "Relationship" had a large effect size (r=-0.52). See Table 4.5.2 for details. The mean rank of the 'Relationship' scores was higher in the ASD group than in the TD group at 24 month (Mean rank: 13.05 vs. 7.95), indicating that children in the ASD group had more problems engaging in a relationship with others than children in the TD group.

Table 4.5.2: Mann-Whitney U test for group comparison of individual m-ADBB item scores

	Facial Expression	Eye Contact	Vocalisation	Activity	Relationship
~12 month					
Mann-Whitney U	34.50	35.00	30.00	50.00	35.00
Z	-1.55	-1.53	-1.94	0.00	-1.51
p-value (1-tailed)*	0.12	0.22	0.12	0.50	0.14
Effect size r	-0.35	-0.34	-0.43	0.00	-0.34
~24 month					
Mann-Whitney U	48.50	34.50	45.00	50.00	24.50
Z	-0.14	-1.55	-0.61	0.00	-2.30
p-value (1-tailed)*	0.46	0.12	0.37	0.50	0.03
Effect size r	-0.03	-0.35	-0.14	0.00	-0.52

*NB: After Bonferroni correction p-value <0.01 is required for statistical significance

4.6 Performance of the m-ADBB in detecting children with ASD

To investigate how useful the m-ADBB is in detecting children with ASD, the m-ADBB classification of whether a child showed problems (i.e. behaviours indicative of social withdrawal) were compared to the true group membership of the children.

4.6.1 Using the m-ADBB cut-off

Because of the small sample size and the possible sampling bias inherent to it, findings will not be extrapolated to the larger population. Therefore, confidence intervals are not being reported. Instead the data that pertains just to this sample will be considered.

Table 4.6.1 shows the m-ADBB classifications for both ASD and TD for 12 months and 24 months in comparison to the children's true group membership. At 12 months, the m-ADBB ratings overall classified 14 children correctly into the ASD or TD groups; this included five of the ten children with ASD and nine of ten children with typically development. At 24 months, the m-ADBB ratings classified 13 children correctly: four of ten children with ASD, and nine of ten children with typical development.

Table 4.6.1 Detection rates based on m-ADBB ratings at both ages

m-ADBB ratings	Group membership		
	ASD	TD	Total
12 months			
Problems	5	1	6
No problems	5	9	14
Total	10	10	20
24 months			
Problems	4	1	5
No problems	6	9	15
Total	10	10	20

Next, the m-ADBB sum scores of each child at both ages were examined. These are shown in figure 4.5.1; the horizontal dotted line represents the cut-off score of 2 or more on the m-ADBB; scores of 2 or more indicate the threshold for concerns regarding social withdrawal behaviour and developmental difficulties. As can be seen from Figure 4.6.1, two typically developing children scored above the cut-off on the m-ADBB at either age (TD9, TD11). The parents of these two children reported no impairments in hearing or vision for

their children at both ages nor did they report any developmental concerns; both rated their children's health as very good in the first two years and reported no parental mood difficulties during that time either, indicating these to be false positive screens. See Appendix C for detailed participant responses.

Within the ASD group, three children scored at or above the cut-off at both ages (ASD1, ASD6 and ASD11); another two children scored at or above the cut-of at 12 months but below it at 24 months (ASD2, ASD5); one child scored below the cut-off at 12 months but above the cut-off at 24 months (ASD7). Three children scored below the cut-off at both ages (ASD8, ASD9 and ASD10); of these, two were diagnosed comparatively late at 5 years 9 months (ASD9) and 10 years (ASD10) respectively. One of these children had received a diagnosis of Asperger's (ASD9), and the parents reported first noticing signs at 30 months, thus later than when the analysed video recordings were made. For another two children videos were available for only one age at which they both scored below the m-ADBB cut-off (ASD3, ASD4). Only one of the four children diagnosed before age 36 months (ASD2, ASD3, ASD4, ASD8) was detected by the m-ADBB at either age. Of the three children diagnosed after age 5 years (ASD9, AD10, ASD11) one was detected with the m-ADBB at both ages, while the others were not detected at either age. For two of the four children whose parents noticed signs before 12 months, the m-ADBB ratings also indicated the presence of problems in the 12 months videos (ASD1, ASD11). Of the six children whose parents had reported loss of skills, two children were correctly identified by the m-ADBB at 12 months (ASD2, ASD5), before the reported loss of skills. Overall, when using the m-ADBB cut-off for social withdrawal as a criterion for a positive screen for ASD, six of the 11 children with ASD were detected at either age 12 months or 24 months.
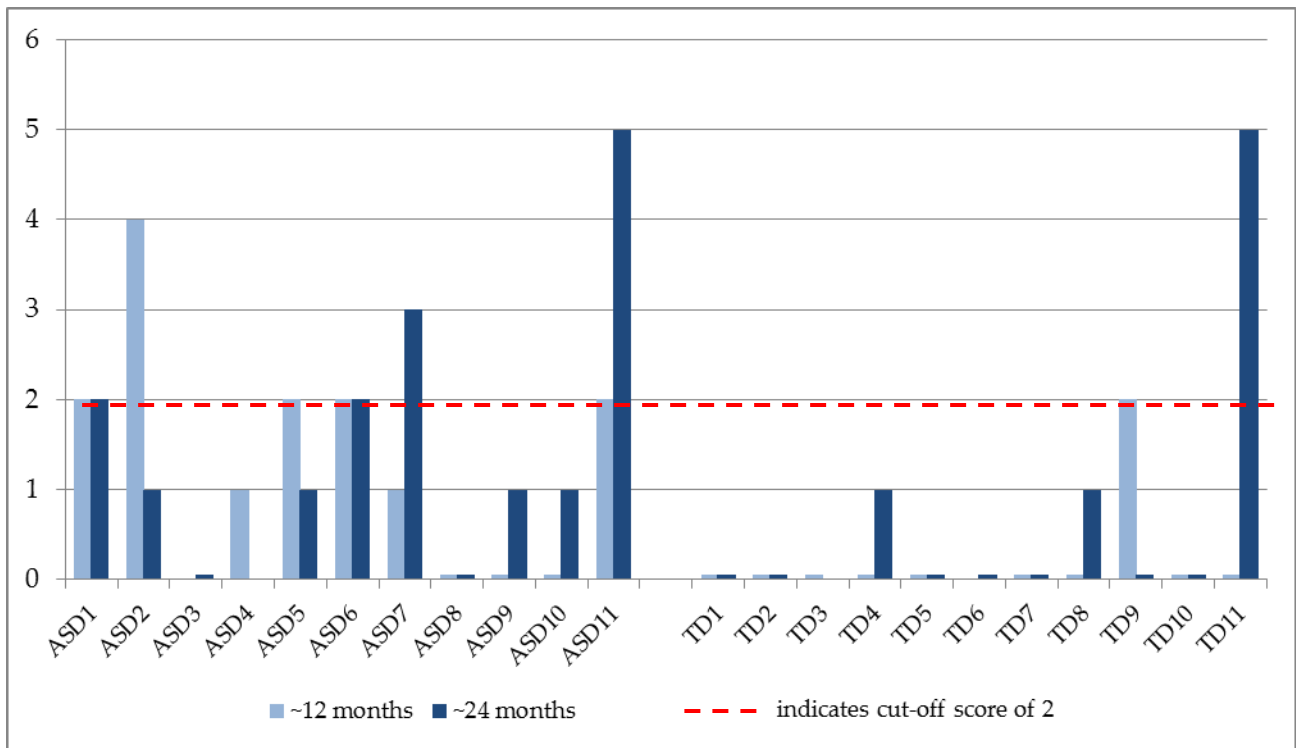
Figure 4.6.1: m-ADBB sum scores for 12 month and 24 month videos for each child
ASD indicates child belongs to ASD group, TD indicated child belongs to TD group

If the cut-off indicating risk on the m-ADBB were to be lowered to a total score of 1 or more (i.e. signifying the presence of at least one possible problem), the detection rate at both ages would improve without adding many false positives. A cut-off of a total score of 1 or more would flag seven of the children with ASD at 12 months, and eight at 24 months. Only two children with ASD would not be detected with the lower cut-off score (ASD3, ASD8). At 12 months, one TD child would be falsely flagged to be at risk (TD9), and at 24 months this would increase to three children (TD4, TD8, TD11) when applying the lower cut-off score.

4.6.2 Using clinical judgement

Clinical judgement was used by both investigators in addition to the m-ADBB scores. As previously explained (see p. 61) clinical judgement was based on observations that were deemed by each investigator as being clinically relevant for early manifestations of ASD independently to whether these were captured by the m-ADBB or not. For example, vocalisations were noted by both investigators to sometimes be qualitatively different in the

24 month videos for children in the ASD group; children from the ASD group were observed to only produce baby talk rather than single words or word approximations, as would be developmentally appropriate for this age.

Both investigators' clinical judgements of group membership at both ages were compared to the actual group membership to determine detection rates. Percentage of agreement for the clinical judgements between both investigators at 12 months was 75% and the IRR was fair (ICC = 0.46); at 24 months 80% of ratings were identical between both investigators, achieving an ICC of 0.76, indicating excellent inter-rater agreement. Because the clinical judgements were not based on standardised coding procedures, 'consensus' agreements were not established. However, for the purpose of combining both investigators' judgements, a child was considered to have ASD if one investigator's judgement indicated this. The clinical judgements of each investigator and the combined judgements for 12 months and 24 months are compared to the children's true membership in Table 4.6.2.

Table 4.6.2 Detection rates based on clinical judgement

	Group membership								
	ASD	TD	Total	ASD	TD	Total	ASD	TD	Total
Clinical judgement	Investigator 1 (SM)			Investigator 2 (BC) *			Both investigators**		
12 months									
ASD	3	1	4	4	1	5	5	2	7
TD	7	9	16	6	9	15	5	8	13
Total	10	10	20	10	10	20	10	10	20
24 months									
ASD	8	2	10	7	1	8	8	3	11
TD	2	8	10	3	9	12	2	7	9
Total	10	10	20	10	10	20	10	10	20

*Investigator 2 had prior knowledge of group status for 5 children

** If at least one investigator's clinical judgement indicated ASD, the child was classified as having ASD

Three and four out of ten children with ASD respectively were correctly identified in the 12 months videos when the researchers used their clinical judgement; taking both researchers' clinical judgements together, the same five children with ASD were identified as when the m-ADBB cut-off of 2 or more was used. Two typically developing children (TD3 and TD9) were falsely classified as having ASD in the 12 months videos when taking both researchers' clinical judgements together. At 24 month, clinical judgement identified seven and eight out of 10 children with ASD respectively. Thus overall four additional children with

ASD were identified in the 24 month videos (ASD2, ASD3, ASD5, ASD10) when clinical judgement of both investigators was combined compared to when the m-ADBB cut-off was used.

Examining individual children, the same five children with ASD were missed in the 12 months videos when investigators used clinical judgement as when the m-ADBB cut-off for social withdrawal was applied (ASD4, ASD7, ASD8, ASD9, ASD10). Two of the six ASD cases that were not detected at 24 months with m-ADBB alone, were also missed when investigators applied clinical judgement (ASD8. ASD9). The typically developing child who was falsely classified as having ASD at 12 months (TD9) with the m-ADBB was also falsely classified as having ASD by the clinical judgement of one researcher (SM).

4.7   Summary of Findings

Children in the ASD group scored statistically significantly higher on the m-ADBB than children in the TD group, in both the 12 months and 24 months videos. There were no statistically significant differences in the individual item scores between the ASD and TD group at both ages, after adjusting for multiple testing; the group difference for the item 'Relationship' at 24 months, however, had a large effect size, indicating children with ASD having more problems engaging in a relationship with others than TD children.

Using the m-ADBB cut-off score of two or more as criterion for a positive ASD screen, five out of ten children with ASD were detected at 12 months and four out of ten children at 24 month; overall six of the 11 children in the ASD group were detected at least at one of the time points (see Figure 4.5.1). Only two of the 11 children in the TD group scored in this range either at 12 or 24 months. The four children, who were diagnosed with ASD by age 30 months, were expected to be showing signs of ASD in their home videos at least by age 24 months; however, only one of them was detected when using the m-ADBB cut-off score. Contrary to expectations, five of the seven children who were diagnosed at a later age, and who therefore may not have been showing early signs of ASD, were correctly identified when using the m-ADBB cut-off for social withdrawal.

When using clinical judgement the investigators identified three to four out of ten children with ASD correctly at 12 months, and seven to eight out of ten at 24 months. This indicates that clinical judgement detects additional children with ASD at 24 months than the m-ADBB alone.

CHAPTER 5: Discussion

5.1   Rational and aim of study

Strong evidence from home-video studies, high-risk sibling studies and parent report studies exists that early behavioural signs of Autism Spectrum Disorder (ASD) can be observed in the first two years of life; in particular deficits in social interaction and communication behaviours have been observed. Despite the large body of evidence, early detection and diagnosis of ASD remains challenging. The clinical usefulness of existing level 1 screening instruments for ASD (e.g. FYI, M-CHAT-R/F) is limited or has been only insufficiently examined (e.g. PEDS, BTISEA). Furthermore, most screening instruments rely on parent report rather than direct observation; they further have been shown to produce high numbers of false positive screening results (e.g. M-CHAT), even when used after the age of 18 months, when early signs of ASD are more reliably observed. While the existing gold standard diagnostic assessments (i.e. ADOS; ADI-R) are highly reliable in detecting children with ASD even before the age of two years, they are not suitable for routine screening due to their time-intensive administration and extensive training requirements. Thus, no reliable and easy to implement instrument is currently available to clinicians to routinely screen infants and toddlers for ASD.

Some of the early ASD related deficits in social communication and interaction behaviours are also characteristic of sustained social withdrawal in infants and toddlers. To date, only one study has investigated whether social withdrawal in infants and toddlers is indicative of ASD, by assessing social withdrawal with the observational ADBB scale (Wendland et al., 2010). Given the encouraging results from the ADBB study, it was considered of merit to investigate the clinical utility of the modified ADBB (m-ADBB) for detecting possible presence of ASD in infants and toddlers. Using the m-ADBB, the current study, analysed home-video recordings of 11 children with ASD and 11 children with typical development from when they were around 12 months and 24 months old. The m-ADBB assesses problems in five social communication and interaction behaviours: facial expression, eye contact, vocalisations, activity level and engaging in relationships with others. It was hypothesised that children with ASD will score statistically significantly higher on the m-ADBB than typically developing children, indicating the presence of more problems. It was considered that this difference would only be clinically useful if it had a large effect size and if the detection rate was meeting the a-priori clinical utility considerations made by the investigators as discussed in Chapter 3. It was further examined whether children with ASD showed a particular pattern of problems across the five behaviours assessed by the m-ADBB.

5.2   Detection of ASD in the current sample

In the current sample, the ASD group scored statistically significantly higher on the m-ADBB at both in their 12 month and 24 month video recordings than the TD group, with large effect sizes at both ages; this lends support to the research hypothesis and indicates the potential ability of the m-ADBB to differentiate between infants and toddlers with ASD and TD.

The previously established cut-off score for social withdrawal on the m-ADBB was applied to indicate possible presence of ASD; this cut-off indicates the presence of at least one definite problem or two possible problems across the five behaviours that the m-ADBB assesses and translates into a total score of two or more. When this cut-off score was applied to the current sample, five out of ten children with ASD were detected in the 12 month videos and four out of ten children with ASD in the 24 month videos. Only one TD child was incorrectly flagged as having ASD at 12 month, and another TD child at 24 month, indicating very low rates of false positive screens. This means observing only five behaviours around age 12 months and 24 months was sufficient to detect up to half of the children with ASD while producing a very low number of false positive results.

These findings are comparable to those reported by Wendland et al. (2010), who used the original eight-item ADBB scale for the analysis of home-videos of children aged six to 18 month. The ADBB was found to identify half of the children with ASD between the ages of six and 12 month as well as between the ages of 12 to 18 month, while producing no false positive screens. The results from the ADBB study by Wendland and colleagues taken together with the results from the current study suggest that screening for social withdrawal behaviours using these scales (ADBB; m-ADBB) may be clinically useful for the early detection of children with ASD.

The findings also indicated that the detection rate of the m-ADBB was further improved when the cut-off score was lowered, so that the presence of only one possible problem or more indicates possible presence of ASD. When this lower cut-off was applied to the current sample, the m-ADBB detected the majority of children with ASD at both time points (seven and eight out of ten respectively at 12 and 24 months). The false positive rate would have remained unchanged at 12 months, and would have increased to three out of 10 at 24 months.

In addition to the strict m-ADBB cut-off criterion, both investigators used their clinical judgement to rate their video observations. These judgements were based on observations that the investigators made in addition to what they had observed using the strict m-ADBB criteria, and which they had deemed clinically relevant of early ASD manifestations. For example, a child might have not scored at or above the m-ADBB cut-off score for

developmental risk, but investigators may have made additional observations leading them to think that the child was indeed in the ASD group rather than the TD group; such as the child showing repetitive behaviour during the video-clip, which is not a behaviour assessed on the m-ADBB. Using clinical judgement in addition to m-ADBB ratings improved detection rates at 24 month but not 12 month compared to using the m-ADBB cut-off score of two or more. Clinical judgement did not detect more children with ASD at 12 and 24 months when compared to using the lower m-ADBB cut-off score.

Previous research has shown that early ASD signs increase in number and severity during the second year of life (Saint-Georges et al., 2010). Thus it was expected that the m-ADBB would detect more children at age 24 months than at age 12 months. However in the current sample the number of children with ASD who were correctly identified was similar at both ages (five at 12 months and four at 24 months), and thus it was not in line with what the literature suggests. This may be due to the small sample size.

Taking a closer look at the children in the current sample who were not flagged by the m-ADBB as having at least one definite or two possible problems across the five assessed behaviours, it appears that an early age of diagnosis (i.e. <36 months) was not linked to a positive screen on the m-ADBB. Indeed, three of the four children diagnosed in the ASD before the age of three years would not have been picked up with the m-ADBB at age 12 month or 24 month (ASD3, ASD4 and ASD8); however, only one of these (ASD3) would have been missed if the lower cut-off score would have been applied. Early signs that the parents of child ASD4 and child ASD8 reported included prolonged screaming, feeding problems, sleeping problem and loss of previously acquired words, problems which are not captured by the ASD. The parents of child ASD3 however, had reportedly noticed a lack of eye contact, absence of smiling, pointing and response to name from age six months onwards. A video recording for age 24 month only was available for this child, and these problems were not noticed in that video clip, suggesting that these problems either had subsided by that age or were not captured in the video. Another child (ASD9) who was screened as having no problems on the m-ADBB at 12 and 24 month did not show any early signs until age 30 months according to parent report. This means, children were not picked up by the m-ADBB if they showed non-ASD specific problems at a young age, or if observable problems did not manifest until after age 24 months.

As described in Chapter 3 Methods, the study investigators deliberated about what they would perceive as a clinically useful outcome in regard to detection rates and false positive rates for the current sample. The rational for this was that statistically significance is mainly affected by sample size, and therefore not sufficient to indicate clinical usefulness, given the

small sample of the current study. For a sample consisting of 11 ASD and 11 TD children, the researchers decided that the m-ADBB must detect three or more children with ASD and no false positives to be clinically useful; or if the detection rate increased to six or seven out of 11 children a maximum of five false positive screens were deemed acceptable. The results of the current study have clearly passed the usefulness criteria of both investigators. While these were subjective criteria, the fact that the results fulfilled these criteria has convinced both researchers to not dismiss the m-ADBB as a potential screener for ASD presence in infants and toddlers, and that further research into its use as screener for early signs of ASD may be merited. Given these results, it further would be useful to survey other health professionals (paediatricians, psychiatrists, early childhood nurses etc.) about what they would perceive as clinically useful detection rates and false positive rates in regard to screening for ASD, and to compare these with the current findings. However this was outside the scope of the current study.

In summary, using the original cut-off score, the m-ADBB identified about half of the children in the ASD group correctly at both ages and at the same time produced a low number of false positive screening results in the TD group at both ages. This means a positive screening result on the m-ADBB strongly suggests that an infant or toddler is may have ASD, while a negative screening result may not rule out ASD. This finding resonates with that for the SACS screener, the only other validated observational level 1 ASD screener, where the absence or abnormality of some behaviour strongly indicated ASD, while the presence or intactness of the same behaviour did not rule out ASD (Barbaro & Dissanayake, 2013).

5.3   Individual m-ADBB items

While the ASD and TD groups differed statistically significantly on the total m-ADBB score, they did not statistically differ on individual m-ADBB item scores after adjusting for multiple testing, potentially due to the small sample size. However, it is noteworthy that the score difference on the 'Relationship' item at 24 months between both groups had a large effect size (r > 0.5). This suggests that at age 24 months children with ASD differed from TD children in their ability to relate to another person. Relating to others requires a combination of behaviours, such as facial expression, eye contact, and vocalisation in order to engage with another person. The difference in 'Relationship' scores together with the lack of difference on the less complex behavioural items suggests that children with ASD can meet criteria for satisfactory eye contact, facial expression and vocalisation, but that they may have difficulties combining these individual behaviours to effectively relate to another person, i.e. make eye contact and smile, gesture or vocalise to another person. This resonates with the research examining the SACS screener, which found that deficits in any single behaviour were found

not to be predictive of ASD, but rather the combination of several deficits (Barbaro & Dissanayake, 2013). Furthermore, evidence for impairments in combining gaze, gesture and vocalisation comes from a variety of high risk siblings and home-video studies (Bernard et al., 2005; Brisson et al., 2011a; Yoder, Stone, Walden, & Malesa, 2009). Initiation of joint attention, which requires a combination of these same behaviours, has also been found to be impaired or reduced in children with ASD (Jones et al., 2014; Rogers, 2009; Saint-Georges et al., 2010; Watson et al., 2013), which again suggests that deficits in the ability to combine behaviours are found more consistently in children with ASD than impairments in the ability to produce a single communication behaviour (e.g. eye contact).

The lack of differences between ASD and TD groups across the remaining m-ADBB items, such as eye contact and facial expression, is not entirely unexpected; the evidence from high-risk sibling studies and home-video studies for deficits in eye contact is inconsistent as discussed in Chapter 1 and Chapter 2; the existing evidence further suggests that reduced quality of eye contact may be more indicative of ASD rather than reduced frequency (Clifford & Dissanayake, 2008; Clifford et al., 2007; Zappella et al., 2015; Zwaigenbaum et al., 2005). As the m-ADBB measure length and frequency of eye contact rather than quality any qualitative differences between ASD and TD groups may have been missed. Again a modification of the m-ADBB scale to include ratings about the quality of the eye contact may improve its usefulness for ASD screening, and this would be worthwhile investigating.

Vocalisations were noted by both investigators to sometimes be qualitatively different in the 24 month videos for children in the ASD group; these were observed to be often producing baby talk rather than single words or word approximations, as would be developmentally appropriate for this age. Again, the m-ADBB only assesses frequency and duration of vocalisations but not their quality, and accounting for quality may enhance the m-ADBB's clinical usefulness in detecting ASD at 24 months. The majority of children in both the ASD and TD group scored satisfactory on the m-ADBB item 'Activity' at both ages; this item therefore appeared to be least able to differentiate between ASD and TD children.

In summary, there were no statistically significant group difference on individual m-ADBB items, however the more complex 'Relationship" item seem to have potential to differentiate between children with ASD and TD; the item 'Activity" appears to be of least value to the detection of ASD risk. Further modifications of the m-ADBB to include quality ratings for eye-contact and vocalisation should be explored.

5.4   Comparison of the m-ADBB to the SACS screener

As discussed in Chapter 1, only one other observational ASD screening instrument for infants and toddlers has been developed to date, which is the screener used in the prospective SACS screening study (Barbaro & Dissanayake, 2010, 2013). The screening items of the SACS screener are tailored to the specific age at assessment, and as such they vary from between 12 to 15 items for the ages 12, 18 and 24 months. While there is some overlap in the behaviours assessed by the SACS with those assessed by the m-ADBB (i.e. eye contact, social smile), the SACS assesses a range of additional behaviours such as pointing, following a point, response to name, showing, producing five to 10 words, waving, imitation and pretend play. These behaviours would not necessarily emerge spontaneously within a clinical routine examination but require prompting and may therefore be more difficult to assess. Some of these additional behaviours, however, have been shown to be useful in differentiating children with ASD from TD children: At 12 months the SACS items that were found to best distinguish between ASD and TD children included deficits in pointing, waving, imitation, eye contact, and response to name; at 18 months these behaviours included deficits in pointing, eye contact, and showing; and at 24 month behaviours that detected children with ASD included deficits in pointing, eye contact, showing, pretend play and waving and these behaviours detected 96.2% of children out of a sample of 50 children with ASD (Barbaro & Dissanayake, 2013). Out of the behaviours that the SACS screener assesses, only eye contact is also assessed by the m-ADBB; thus, it is noteworthy that the m-ADBB detected up to six out of 10 children with ASD in the current sample across both time points, suggesting that the other social withdrawal behaviours assessed by the m-ADBB appear to be also useful in detecting possible presence of ASD.

5.5   Theoretical implications

The systematic review conducted as part of this study found that evidence from home-video studies of early behavioural signs of ASD is plentiful but only consistent for a few early signs; these were: response to name, looking at others, facial expression, affect quality, eye contact, and use of gestures. Variations in findings across studies of both may have been caused by differences in sample characteristics and type of observational approach employed by these studies. Since most studies used small sample sizes, replication in larger study samples is merited to establish robust evidence for several other behaviours suggested by these studies as potential early signs of ASD. Despite the variability in study design and methods, the systematic review further found consistent findings for some behaviours (as listed above) across different analysis approaches, namely time-sampled coding and clinical rating scales, suggesting that less detailed observations (e.g. with clinical rating scales) can be as useful in

detecting early signs of ASD as the more detailed time-sampled coding grids. This also means that early signs of ASD can be observed in non-standardised settings, such as varying clinical settings. Together this means that less detailed observational ratings scales may be of merit for ASD screening in routine clinical practice.

The findings of the current home-video study also add to the literature about the link between social withdrawal and early signs of ASD, which has previously been investigated by only one study (Wendland et al., 2010). The current study results confirmed previous research findings in that some children with ASD show more social withdrawal in infancy and toddlerhood than those with typical development as assessed by the ADBB and m-ADBB.

5.6 Clinical implications

Examining a validated screening instrument for social withdrawal in infants and toddlers, the current study found that the observation of only five behaviours was sufficient to detect up to half of the children with ASD in the current sample. Together with the findings from the previous home-video study examining the original eight-item ADBB (Wendland et al., 2010) the current findings show that observation of a few social withdrawal behaviours has the potential to be clinically useful in detecting infants and toddlers that may benefit from further assessments for ASD. This means a positive screen on the m-ADBB indicates not only developmental risk in general but also suggests possible presences of ASD, and thus indicates the need for further investigations regarding ASD.

Additionally, the briefness and ease of use of the m-ADBB is noteworthy when considering clinical usefulness; recent research efforts have turned towards more complex and technologically aided screening approaches, such as eye tracking and neuroimaging, to detect ASD in infants (Jones et al., 2014; Jones & Klin, 2013). These sophisticated technological approaches however, cannot be easily translated to routine screening as they require expensive technology and are time-intensive in measuring and analysing assessment data. Thus, the current findings are encouraging for clinical practice, as they suggest that a brief observation, such as with the m-ADBB, may help with the early detection of ASD; the m-ADBB could be used alongside other existing screening instruments, such as parent measures (e.g. M-CHAT-R/F). Also modifications to some m-ADBB items may improve its usefulness further.

5.7   Strengths and limitations

The same limitations apply to this study as to other home-video studies. Firstly video-material was not standardised in its content and lengths. However, the average length of video footage was similar in both groups; furthermore, the videos were similar in the type of settings in which the children were video recorded, mainly playing at home or family routines. Selection of videos did not appear to have been biased to show particularly typical autistic behaviours or particularly good examples of interactions according to parent report. Having assessed for these potential sources of bias is a strength of the current study, and so is the fact that videos were comparable in length and setting between both groups. Secondly, the majority of video recordings did not meet the minimum length of three minutes originally set for inclusion. The briefness of most of the video recordings may have affected the validity of the researchers' observations as the behaviour captured in those videos may not have been representative of the child's usual behaviour.

Another limitation is that the diagnostic status of children was not ascertained with gold-standard diagnostic assessments, such as the ADI-R and ADOS (Lord et al., 2000; Rutter et al., 2003), by the investigators themselves; rather it was based on parent report of previous diagnostic assessments conducted by health professionals. Conducting diagnostic assessments was not deemed feasible within the scope of the current study, where recruitment efforts expanded to other states and countries. To ascertain that control participants were true controls, parents were asked to rate their children's current social responsiveness on a validated screener for ASD for the ages 4 to 18 years, the SRS-Brief (Moul et al., 2014). The ratings for all children in the ASD group fell above the cut-off score indicative for ASD, while one of the ratings of the children in the TD group fell above that cut-off score. This information lends support to the validity of the parent's report of their children's ASD diagnosis or typically development respectively. The author of this thesis acknowledges that the validity of the SRS-brief ratings for some children may have been limited who were younger than four years old at the time of recruitment, the lower age limit for which the SRS-brief has been validated.

The current study sample was small which limits representativeness of the findings due to sampling bias; also most parents in the current study sample had university level education, indicating that they were not a representative sample of the general population. It is noteworthy that four children in the ASD group were diagnosed before the age of three, which is earlier than the average age of diagnosis of four years found in Australia (Young et al., 2003); this again may be reflective of the high educational level of the participating parents. Fountain, King, and Bearman (2011) found that children of highly educated parents are diagnosed earlier than children of parents with lower educational achievements.

However, a sufficiently large enough sample size was achieved to obtain a statistically significant difference in the group comparison of total m-ADBB scores. The small sample size was mostly due to difficulties in recruitment as outlined in Chapter 1. Nevertheless, the sample size was comparable to that of previous home-video studies.

To remediate the recruitment difficulties, several clinicians in the field were approached by the author of this thesis; these clinicians expressed willingness to disseminate information about the study to their clients (using the approved Participant Information Statements etc.). This appeared to be a feasible solution to the recruitment problem, especially as these clients were used to providing video footage of their child during consultation, as reported by the clinicians. However, unexpectedly, the University of Sydney Human Research Ethics Committee (HREC) refused approval of this recruitment strategy. They argued that this was coercive and that discussing early signs of ASD had no place in a consultation, even if this involved assessment of ASD. But the clinicians who had been approached, stated that it was standard practice for parents to bring video recordings of their children to consultations, to better illustrate their children's behavioural problems, as these do not necessarily manifest during the course of a consultation. Through personal communication with two clinical psychologists working with families of children with ASD, the author of this thesis learned that parents wished to be partners in the care for their children with ASD, and clinicians therefore preferred to have a reciprocal relationship with the parents; and this could include parents participating in research to further the understanding of their children's condition. Thus, it appears that the HREC may not have been aware of standard clinical practice and the nature of the relationship between the treating clinician and the child's parents when rejecting this recruitment approach. Furthermore, one clinician and academic, with whom the author discussed the recruitment approach as well as the HREC's response, commented that the stance of the HREC risks fragilising families of children with developmental difficulties as not being capable of making informed decisions, ignoring that these families tend to be very skilled in negotiating the health system and in securing support for their children, which required a high level of assertiveness and decision making abilities. After ten weeks of negotiations the HREC finally approved of this recruitment approach without providing an explanation why the approach, previously deemed unethical, was now perceived to be ethical. By this time the initial momentum of the involved clinicians for recruitment was lost and this approach did no yield further participants.

Another limitation is that only fair to good inter-rater reliability was achieved for the 12 months ratings which may have limited the validity of the 12 months ratings. However, for all discrepant ratings, consensus was reached between both investigators, and only those

consensus ratings have been used in all further analysis; this should have limited the impact of the bias arising from less than excellent rater concordance on study results.

Lastly, the author of this thesis was not blind to the group membership of five participants, as some participants were recruited from her own social network, and thus, these children were known to her. However, as shown in Chapter 4 this did not impact on concordance between ratings of both researchers, and thus, while it can be seen as limitation, it did not lead to any measurable bias.

5.8   Future directions

Because the current study was exploratory in nature, a replication of the results is merited, in particular with a larger sample to establish the best cut-off score of the m-ADBB for maximum sensitivity and specificity. Also, as noted before, a survey of other health professionals working in early childhood care would be useful to establish consensus on detection rates and false positive rates that could be considered useful in regard to level 1 screening for ASD. Potential modifications to the m-ADBB item, such as inclusion of quality ratings of eye contact and vocalisation, could also be further explored. And lastly, it would be of interest to examine the performance of the m-ADBB and the SACS screener within the same sample of home-videos to establish the comparative merit of each screener at each age.

5.9   Conclusions

The results of the current study suggest that children diagnosed with ASD show more social withdrawal on the m-ADBB in the first two years of life than typical developing children. The results further indicate that the observation of just five social communication and interaction behaviours may be clinically useful for the detection of infants and toddlers with possible presence of ASD and who therefore would benefit from further assessment. Further research in a larger sample is required to establish the m-ADBB's sensitivity and specificity for detection of ASD and to examine how modifications of the m-ADBB may improve its utility for screening.

References

Adrien, J., Barthelemy, C., Perrot, A., Roux, S., Lenoir, P., Hameury, L., & Sauvage, D. (1992a). Validity and reliability of the Infant Behavioral Summarized Evaluation (IBSE): A rating scale for the assessment of young children with autism and developmental disorders. *Journal of Autism and Developmental Disorders, 22*(3), 375-394. doi: http://dx.doi.org/10.1007/BF01048241

Adrien, J., Gattegno, M. P., Streri, A., Reynaud, L., & Barthelemy, C. (2005). Early development and creativity in autistic children. *Archives de Pediatrie, 12*(6), 858-860.

Adrien, J. L., Lenoir, P., Martineau, J., Perrot, A., Hameury, L., Larmande, C., & Sauvage, D. (1993). Blind ratings of early symptoms of autism based upon family home movies. *J Am Acad Child Adolesc Psychiatry, 32*(3), 617-626. doi: 10.1097/00004583-199305000-00019

Adrien, J., Perrot, A., Hameury, L., Martineau, J., Roux, S., & Sauvage, D. (1991). Family home movies: Identification of early autistic signs in infants later diagnosed as autistics. *Brain Dysfunction, 4*(6), 355-362.

Adrien, J., Perrot, A., Sauvage, D., Leddet, I., Larmande, C., Hameury, L., & Barthelemy, C. (1992b). Early symptoms in autism from family home movies. Evaluation and comparison between 1st and 2nd year of life using I.B.S.E. scale. *Acta Paedopsychiatr, 55*(2), 71-75.

Adrien, J., Roux, S., Couturier, G., Malvy, J., Guerin, P., Debuly, S., . . . Barthelemy, C. (2001). Towards a new functional assessment of autistic dysfunction in children with developmental disorders: the Behaviour Function Inventory. *Autism, 5*(3), 249-264.

APA. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text rev.)* American Psychiatric Association (Ed.)   doi:10.1176/appi.books.9780890423349.

APA. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.* American Psychiatric Association (Ed.)   Retrieved from dsm.psychiatryonline.org

Baranek, G. T. (1999). Autism during infancy: a retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age. *J Autism Dev Disord, 29*(3), 213-224.

Baranek, G. T. . (1997). Early predictors of autism using retrospective video analysis. *Dissertation Abstracts International: Section B: The Sciences and Engineering., 57*(11-B), 7247.

Baranek, G. T., Barnett, C. R., Adams, E. M., Wolcott, N. A., Watson, L. R., & Crais, E. R. (2005). Object play in infants with autism: methodological issues in retrospective video analysis. *Am J Occup Ther, 59*(1), 20-30.

Barbaro, J., & Dissanayake, C. (2009). Autism spectrum disorders in infancy and toddlerhood: a review of the evidence on early signs, early identification tools, and early diagnosis. *Journal of Developmental & Behavioral Pediatrics, 30*(5), 447-459.

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*(5), 376-385.

Barbaro, J., & Dissanayake, C. (2013). Early markers of autism spectrum disorders in infants and toddlers prospectively identified in the Social Attention and Communication Study. *Autism, 17*(1), 64-86.

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*(4), 334-347. doi: http://dx.doi.org/10.1016/j.pedn.2010.04.007

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

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

Bayrami, L. (2011). Early identification of autistic spectrum disorders: A retrospective analysis of early social-emotional and communicative indicators. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 71*(9-B), 5823.

Begeer, S., Mandell, D., Wijnker-Holmes, B., Venderbosch, S., Rem, D., Stekelenburg, F., & Koot, H. M. (2013). Sex differences in the timing of identification among children and adults with autism spectrum disorders. *J Autism Dev Disord, 43*(5), 1151-1156. doi: 10.1007/s10803-012-1656-z

Bent, C. A., Dissanayake, C., & Barbaro, J. (2015). Mapping the diagnosis of autism spectrum disorders in children aged under 7 years in Australia, 2010-2012. *Med J Aust, 202*(6), 317-320.

Bernabei, P., Camaigni, L., & Levi, G. (1998). An Evaluation of Early Development in Children with Autism and Pervasive Developmental Disorders from Home Movies: Preliminary Findings. *Autism, 2*(3), 243-258. doi: 10.1177/1362361398023003

Bernard, J., Adrien, J., Roux, S., & Barthelemy, C. (2005). A home video analysis of the relational complex gesture - gaze - oralization in children with autism aged 12 to 36 months. *Devenir, 17*(2), 105-122. doi: http://dx.doi.org/10.3917/dev.052.0105

Bernard, J., Adrien, J., Roux, S., & Barthelemy, C. (2006). Evolution of frequency of gestures from 1 to 3 years of age in boys with autism: A home movie study. *Devenir, 18*(3), 245-261.

Book, L. A. (2010). Early red flags for autism spectrum disorders in toddlers in the home environment. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 71*(3-B), 1662.

Briggs-Gowan, M. J. , Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. . (2004). The Brief Infant-Toddler Social and Emotional Assessment: Screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology, 29*(2), 143-155.

Brisson, J., Serres, J., & Adrien, J. L. (2011a). Early signs of dysfunction in infants (from birth to 6 months of age) subsequently diagnosed as autistic and parental soliciting behaviour. What about early interaction? *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant, 23*(115), 427-432.

Brisson, J., Warreyn, P., Serres, J., Foussier, S., & Adrien-Louis, J. (2011b). Motor anticipation failure in infants with autism: a retrospective analysis of feeding situations. *Autism, 16*(4), 420-429. doi: 10.1177/1362361311423385

Carroll, L. (2015). Autism Spectrum Disorders: Implication sfrom the changes in DSM-V. *Journal of Mental Health for Children and Adolescents with Intellectual and Developmental Disabilities: An Educational Resource, 6*(1), 12-14.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *JAMA, 285*(24), 3093-3099. doi: 10.1001/jama.285.24.3093

Chawarska, K., Paul, R., Klin, A., Hannigen, S., Dichtel, L. E., & Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism & Developmental Disorders, 37*(1), 62-72.

Chin, K. (2009). Evaluation of early social contingency behavior in children with autism. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 70*(1-B), 682.

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics, 131*(4), e1121-1127.

Clifford, S. M., & Dissanayake, C. (2008). The early development of joint attention in infants with autistic disorder using home video observations and parental interview. *J Autism Dev Disord, 38*(5), 791-805. doi: 10.1007/s10803-007-0444-7

Clifford, S., Young, R., & Williamson, P. (2007). Assessing the early characteristics of autistic disorder using video analysis. *J Autism Dev Disord, 37*(2), 301-313. doi: 10.1007/s10803-006-0160-8

Cohen, J. . (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, New Jersey: LEA.

Colgan, S. E., Lanter, E., McComish, C., Watson, L. R., Crais, E. R., & Baranek, G. T. (2006). Analysis of social interaction gestures in infants with autism. *Child Neuropsychol, 12*(4-5), 307-319. doi: 10.1080/09297040600701360

Constantino, J. N., & Gruber, C. P. (2005). *Social Responsiveness Scale (SRS)*. Los Angeles: Western Psychological Services.

Costanzo, V., Chericoni, N., Amendola, F. A., Casula, L., Muratori, F., Scattoni, M. L., & Apicella, F. (2015). Early detection of autism spectrum disorders: From retrospective home video studies to prospective 'high risk' sibling studies. *Neurosci Biobehav Rev, 55*, 627-635. doi: 10.1016/j.neubiorev.2015.06.006

Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Dev Psychopathol, 20*(03), 775-803. doi: doi:10.1017/S0954579408000370

Dawson, G., Rogers, S. J., Munson, J.A., Smith, M., Winter, J., Greenson, J., . . . Varley, J. (2010). Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model. *Pediatrics, 125*(1), e17-e23. doi: 10.1542/peds.2009-0958

Degenne, C., Serres, J., Gattegno, M. P., & Adrien, J.-L. (2009). Preliminary study of interactions and motricity issues of few-days-old to 6-months-old babies who exhibited an autistic symptom later on: Microscopic analysis based on temporal sampling of family videotapes, using the observer software. *Devenir, 21*(4), 265-294.

Dietz, C., Swinkels, S. H., Daalen, E., 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*(6), 713-722. doi: 10.1007/s10803-006-0114-1

Eapen, V., Črnčec, R., Woolfenden, S., & Blackmore, R. (2013). Screening for Autism Spectrum Disorders Using the PEDS and M-CHAT. *Journal of Mental Health Research in Intellectual Disabilities, 7*(1), 1-13. doi: 10.1080/19315864.2012.704489

Eaves, L. C., Wingert, H., & Ho, H. H. (2006). Screening for autism: Agreement with diagnosis. *Autism, 10*(3), 229-242. doi: 10.1177/1362361306063288

Esposito, G., & Venuti, P. (2008). Persistence of motoric anomalies in subjects with autistic spectrum disorders. *Infanzia e Adolescenza, 7*(3), 138-148.

Esposito, G., & Venuti, P. (2010). Developmental changes in the fundamental frequency (f0) of infants' cries: A study of children with Autism Spectrum Disorder. *Early Child Development and Care, 180*(8), 1093-1102. doi: http://dx.doi.org/10.1080/03004430902775633

Esposito, G., & Venuti, P. (2010a). Understanding early communication signals in autism: a study of the perception of infants' cry. *Journal of Intellectual Disability Research, 54*(3), 216-223. doi: http://dx.doi.org/10.1111/j.1365-2788.2010.01252.x

Esposito, G., Venuti, P., & Bornstein, M. (2011). Assessment of distress in young children: A comparison of autistic disorder, developmental delay, and typical development. *Research in Autism Spectrum Disorders, 5*(4), 1510-1516. doi: http://dx.doi.org/10.1016/j.rasd.2011.02.013

Esposito, G., Venuti, P., Maestro, S., & Muratori, F. (2009). An exploration of symmetry in early autism spectrum disorders: Analysis of lying. *Brain and Development, 31*(2), 131-138. doi: http://dx.doi.org/10.1016/j.braindev.2008.04.005

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

Fountain, C., King, M. D., & Bearman, P. S. (2011). Age of diagnosis for autism: individual and community factors across 10 birth cohorts. *Journal of Epidemiology & Community Health, 65*(6), 503-510.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2-18. doi: 10.1037/a0024338

Gardner, L. M., Murphy, L., Campbell, J. M., Tylavsky, F., Palmer, F. B., & Graff, J. C. (2013). Screening accuracy for risk of autism spectrum disorder using the Brief Infant-Toddler Social and Emotional Assessment (BITSEA). *Research in Autism Spectrum Disorders, 7*(5), 591-600. doi: http://dx.doi.org/10.1016/j.rasd.2013.01.004

Gattegno, M., Reynaud, L., Streri, A., Barthelemy, C., & Adrien, J. (2005). Early interactive, joint attention and motor dysfunctionings in 9 infants aged from 4 to 6 months and furtherly diagnosed as autistics. *A N A E Approche Neuropsychologique des Apprentissages chez l'Enfant, 17*(3-4[83-84]; 83-84), 152-164.

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Goldsmith, H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in

autism. *Journal of Child Psychology and Psychiatry, 49*(1), 43-50. doi: http://dx.doi.org/10.1111/j.1469-7610.2007.01820.x

Gerwing, J. J. (2009). Quantifying infant social responsiveness: Microanalysis of home videos of a set of triplets for early indications of autism. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 70*(11-B), 7266.

Giarelli, E., Wiggins, L. D., Rice, C. E., Levy, S. E., Kirby, R. S., Pinto-Martin, J., & Mandell, D. (2010). Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disabil Health J, 3*(2), 107-116. doi: 10.1016/j.dhjo.2009.07.001

Glascoe, F. P. (2006). *Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children*. Nashville, TN: Ellsworth & Vandermeer Press.

Glascoe, F. P., Macias, M. M., Wegner, L. M., & Robertshaw, N. S. (2007). Can a Broadband Developmental-Behavioral Screening Test Identify Children Likely to Have Autism Spectrum Disorder? *Clinical Pediatrics, 46*(9), 801-805. doi: 10.1177/0009922807303928

Glascoe, F. P., & Squires, J. K. (2009). Letter to the Editor. *Journal of Developmental & Behavioral Pediatrics, 30*(2), 174.

Goldberg, W. A., Thorsen, K. L., Osann, K., & Spence, M. A. (2008). Use of home videotapes to confirm parental reports of regression in autism. *J Autism Dev Disord, 38*(6), 1136-1146. doi: 10.1007/s10803-007-0498-6

Grønborg, T. K., Schendel, D. E., & Parner, E. T. (2013). Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: A population-based cohort study. *JAMA Pediatrics, 167*(10), 947-953. doi: 10.1001/jamapediatrics.2013.2259

Guedeney, A., & Fermanian, J. (2001). A validity and reliability study of assessment and screening for sustained withdrawal reaction in infancy: The Alarm Distress Baby scale. *Infant Mental Health Journal, 22*(5), 559-575. doi: 10.1002/imhj.1018

Guedeney, A., Foucault, C., Bougen, E., Larroque, B., & Mentré, F. (2008). Screening for risk factors of relational withdrawal behaviour in infants aged 14–18 months. *European Psychiatry, 23*(2), 150-155. doi: http://dx.doi.org/10.1016/j.eurpsy.2007.07.008

Guedeney, A., Matthey, S., & Puura, K. (2013). Social withdrawal behavior in infancy: A history of the concept and a review of published studies using the alarm distress baby scale. *Infant Mental Health Journal*, n/a-n/a. doi: 10.1002/imhj.21412

Guidetti, M., Turquois, L., Adrien, J., Barthelemy, C., & Bernard, J. (2004). Pragmatic aspects of communication and language in typical children and in children later diagnosed

with autism. *Psychologie Francaise, 49*(2), 131-144. doi: http://dx.doi.org/10.1016/j.psfr.2004.04.001

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol, 8*(1), 23-34.

Happé, F., & Ronald, A. (2008). The 'Fractionable Autism Triad': A Review of Evidence from Behavioural, Genetic, Cognitive and Neural Research. *Neuropsychology Review, 18*(4), 287-304. doi: http://dx.doi.org/10.1007/s11065-008-9076-8

Hartley, C., Pretorius, K., Mohamed, A., Laughton, B., Madhi, S., Cotton, M. F., . . . Seedat, S. (2010). Maternal postpartum depression and infant social withdrawal among human immunodeficiency virus (HIV) positive mother-infant dyads. *Psychol Health Med, 15*(3), 278-287. doi: 10.1080/13548501003615258

Howlin, P. (1997). Prognosis in autism: do specialist treatments affect long-term outcome? *European Child & Adolescent Psychiatry, 6*(2), 55-72. doi: 10.1007/bf00566668

IBM Corp. (Released 2013). IBM SPSS Statistics for Windows, Version 22 (Version 22). Armonk, NY: IBM Corp.

Jones, E. J., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: a review of prospective studies of infants at risk. *Neuroscience & Biobehavioral Reviews, 39*, 1-33.

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

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., . . . Fein, D. (2008). The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism & Developmental Disorders, 38*(5), 827-839.

Kobayashi, R., & Murata, T. (1998). Setback phenomenon in autism and longterm prognosis. *Acta Psychiatrica Scandinavica, 98*(4), 296-303. doi: 10.1111/j.1600-0447.1998.tb10087.x

Lai, M., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *J Am Acad Child Adolesc Psychiatry, 54*(1), 11-24. doi: 10.1016/j.jaac.2014.10.003

Lehmann, E. L. . (1998). *Nonparametrics: Statistical Methods Based on Ranks* New York, United States: Springer.

Levy, S. E., Mandell, D. S., & Schultz, R. T. (2009). Autism. *Lancet, 374*(9701), 1627-1638. doi: 10.1016/s0140-6736(09)61376-3

Lipsey, M. W. (1990). *Design sensitivity*. Newbury Park, California: Sage Publications.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., 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*(3), 205-223. doi: 10.1023/A:1005592401947

Lösche, G. (1990). Sensorimotor and Action Development in Autistic Children from Infancy to Early Childhood. *Journal of Child Psychology and Psychiatry, 31*(5), 749-761. doi: 10.1111/j.1469-7610.1990.tb00815.x

Maestro, S., Casella, C., Milone, A., Muratori, F., & Palacio-Espasa, F. (1999). Study of the onset of autism through home movies. *Psychopathology, 32*(6), 292-300.

Maestro, S., Milone, A., Muratori, F., & Casella, C. (1999). Pervasive developmental disorders: A research into precocious signs. *I disturbi pervasivi dello sviluppo: Una ricerca sui segni precoci, 25*(2), 306-312.

Maestro, S., Muratori, F., Barbieri, F., Casella, C., Cattaneo, V., Cavallaro, M. C., . . . Palacio-Espasa, F. (2001). Early behavioral development in autistic children: the first 2 years of life through home movies. *Psychopathology, 34*(3), 147-152. doi: 49298

Maestro, S., Muratori, F., Cavallaro, M. C., Pecini, C., Cesari, A., Paziente, A., . . . Palacio-Espasa, F. (2005a). How Young Children Treat Objects and People: An Empirical Study of the First Year of Life in Autism. *Child Psychiatry and Human Development, 35*(4), 383-396. doi: 10.1007/s10578-005-2695-x

Maestro, S., Muratori, F., Cavallaro, M. C., Pei, F., Stern, D., Golse, B., & Palacio-Espasa, F. (2002). Attentional skills during the first 6 months of age in autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry, 41*(10), 1239-1245. doi: 10.1097/00004583-200210000-00014

Maestro, S., Muratori, F., Cesari, A., Cavallaro, M. C., Paziente, A., Pecini, C., . . . Sommario, C. (2005b). Course of autism signs in the first year of life. *Psychopathology, 38*(1), 26-31.

Maestro, S., Muratori, F., Cesari, A., Pecini, C., Apicella, F., & Stern, D. (2006). A view to regressive autism through home movies. Is early development really normal? *Acta Psychiatrica Scandinavica, 113*(1), 68-72.

Malvy, J., Adrien, J., Roux, S., Lataste, C., & Sauvage, D. (1994). Infantile autism and home movies: Functional evaluation of first signs. *A.N.A.E. Approche Neuropsychologique des Apprentissages chez l'Enfant, 6*(1), 33-37.

Mars, A. E., Mauk, J. E., & Dowrick, P. W. (1998). Symptoms of pervasive developmental disorders as observed in prediagnostic home videos of infants and toddlers. *J Pediatr, 132*(3 Pt 1), 500-504.

Matson, J. L. (2007). Determining treatment outcome in early intervention programs for autism spectrum disorders: A critical analysis of measurement issues in learning based interventions. *Research in Developmental Disabilities, 28*(2), 207-218. doi: http://dx.doi.org/10.1016/j.ridd.2005.07.006

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

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities, 30*(6), 1107-1114. doi: http://dx.doi.org/10.1016/j.ridd.2009.06.003

Matthey, S., ČrnČec, R., Hales, A., & Guedeney, A. (2013). A description of the modified alarm distress baby scale (m-adbb): An instrument to assess for infant social withdrawal. *Infant Mental Health Journal*. doi: 10.1002/imhj.21407

Matthey, S., Guedeney, A., Starakis, N., & Barnett, B. (2005). Assessing the social behavior of infants: Use of the ADBB Scale and relationship to mother's mood. *Infant Mental Health Journal, 26*(5), 442-458. doi: 10.1002/imhj.20061

Mattila, M. L., Kielinen, M., Linna, S. L., Jussila, K., Ebeling, H., Bloigu, R., . . . Moilanen, I. (2011). Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: an epidemiological study. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*(6), 583-592.e511.

McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry, 51*(4), 368-383. doi: 10.1016/j.jaac.2012.01.007

Mitchell, S., Cardy, J. O., & Zwaigenbaum, L. (2011). Differentiating autism spectrum disorder from other developmental delays in the first two years of life. *Developmental Disabilities Research Reviews, 17*(2), 130-140.

Mitchell, S. J. (2015). Infants at risk for autism spectrum disorder: Gestures in infants and mothers. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 76*(4-B(E)), No Pagination Specified.

Moul, C., Cauchi, A., Hawes, D. J., Brennan, J., & Dadds, M. R. (2014). Differentiating Autism Spectrum Disorder and Overlapping Psychopathology with a Brief Version of the

Social Responsiveness Scale. *Child Psychiatry and Human Development*. doi: 10.1007/s10578-014-0456-4

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: history, application, and current practice. *Clin Psychol Rev, 21*(3), 421-446.

Osterling, J. A. (1998). Early identification of infants with autism and infants with a global developmental delay. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 58*(12-B), 6819.

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Dev Psychopathol, 14*(02), 239-251. doi: doi:10.1017/S0954579402002031

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: a study of first birthday home videotapes. *J Autism Dev Disord, 24*(3), 247-257.

Ozonoff, S., Heung, K., Byrd, R., Hansen, R.L., & Hertz-Picciotto, I. (2008a). The onset of autism: patterns of symptom emergence in the first years of life. *Autism Research, 1*(6), 320-328. doi: 10.1002/aur.53

Ozonoff, S., Iosif, A. M., Young, G. S., Hepburn, S., Thompson, M., Colombi, C., . . . Rogers, S. J. (2011). Onset patterns in autism: correspondence between home video and parent report. *J Am Acad Child Adolesc Psychiatry, 50*(8), 796-806 e791. doi: 10.1016/j.jaac.2011.03.012

Ozonoff, S., Young, G. S., Goldring, S., Greiss-Hess, L., Herrera, A. M., Steele, J., . . . Rogers, S. J. (2008b). Gross motor development, movement abnormalities, and early identification of autism. *Journal of Autism and Developmental Disorders, 38*(4), 644-656.

Palomo, R., Belinchón, M., & Ozonoff, S. (2006). Autism and Family Home Movies: A Comprehensive Review. *Journal of Developmental & Behavioral Pediatrics, 27*(2), S59-S68.

Pandey, J., Verbalis, A. D., Robins, D. L., Boorstein, H. C., 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*(5), 513-535. doi: 10.1177/1362361308094503

Phagava, H., Muratori, F., Einspieler, C., Maestro, S., Apicella, F., Guzzetta, A., . . . Cioni, G. (2008). General movements in infants with autism spectrum disorders. *Georgian Med News*(156), 100-105.

Pinto-Martin, J. A., Young, L. M., Mandell, D. S., Poghosyan, L., Giarelli, E., & Levy, S. E. (2008). Screening strategies for autism spectrum disorders in pediatric primary care. *J Dev Behav Pediatr, 29*(5), 345-350. doi: 10.1097/DBP.0b013e31818914cf

Poon, K. K.-L. (2005). Predicting language and intellectual outcomes of children with autism spectrum disorders based on the retrospective analysis of infant home videos. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 66*(4-A), 1321.

Poon, K., Watson, L. R., Baranek, G., & Poe, M. (2012). To What Extent Do Joint Attention, Imitation, and Object Play Behaviors in Infancy Predict Later Communication and Intellectual Functioning in ASD? *Journal of Autism & Developmental Disorders, 42*(6), 1064-1074. doi: 10.1007/s10803-011-1349-z

Puura, K., Guedeney, A., Mäntymaa, M., & Tamminen, T. (2007). Detecting infants in need: Are complicated measures really necessary? *Infant Mental Health Journal, 28*(4), 409-421. doi: 10.1002/imhj.20144

Puura, K., Mantymaa, M., Luoma, I., Kaukonen, P., Guedeney, A., Salmelin, R., & Tamminen, T. (2010). Infants' social withdrawal symptoms assessed with a direct infant observation method in primary health care. *Infant Behav Dev, 33*(4), 579-588. doi: 10.1016/j.infbeh.2010.07.009

Receveur, C., Lenoir, P., Desombre, H., Roux, S., Barthelemy, C., & Malvy, J. (2005). Interaction and imitation deficits from infancy to 4 years of age in children with autism: a pilot study based on videotapes. *Autism, 9*(1), 69-82. doi: 10.1177/1362361305049030

Reznick, J. S., Baranek, G. , & Watson, L. R. (2005). *Developmental Concerns Questionnaire*. Chapel Hill, NC: University of North Carolina.

Reznick, J. S., Baranek, G.T, Reavis, S., Watson, L. R, & Crais, E. R. (2007). A Parent-Report Instrument for Identifying One-Year-Olds at Risk for an Eventual Diagnosis of Autism: The First Year Inventory. *Journal of Autism and Developmental Disorders, 37*(9), 1691-1710. doi: 10.1007/s10803-006-0303-y

Robins, D. L., Casagrande, K., Barton, M., Chen, C.-M. A., Dumont-Mathieu, T., & Fein, D. (2013). Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). *Pediatrics*. doi: 10.1542/peds.2013-1813

Robins, D. L., & Dumont-Mathieu, T. M. (2006). Early screening for autism spectrum disorders: update on the modified checklist for autism in toddlers and other measures. *Journal of Developmental & Behavioral Pediatrics, 27*(2 Suppl), S111-119.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An Initial Study Investigating the Early Detection of Autism and Pervasive Developmental Disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131-144.

Rogers, S.J. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research, 2*(3), 125-137. doi: 10.1002/aur.81

Rutter, M., LeCouteur, A., & Lord, C. . (2003). *Autism Diagnostic Interview – Revised. .* Los Angeles: Western Psychological Services.

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

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale - 2nd Edition.* Los Angeles: Western Psychological Services.

Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Ment Retard Dev Disabil Res Rev, 10*(4), 234-247. doi: 10.1002/mrdd.20038

Swinkels, S. H., Dietz, C., van Daalen, E., Kerkhof, I. H., 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 & Developmental Disorders, 36*(6), 723-732.

Thorsen, K. L., Goldberg, W. A., Osann, K., & Spence, M. A. (2008). Birthday and non-birthday videotapes: the importance of context for the behavior of young children with autism. *J Autism Dev Disord, 38*(6), 1047-1058. doi: 10.1007/s10803-007-0479-9

Towbin, K. E., Pradella, A., Gorrindo, T., Pine, D. S., & Leibenluft, E. (2005). Autism spectrum traits in children with mood and anxiety disorders. *J Child Adolesc Psychopharmacol, 15*(3), 452-464. doi: 10.1089/cap.2005.15.452

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: research, theory, and clinical applications. *J Child Psychol Psychiatry, 42*(1), 3-48.

Turner-Brown, L. M., Baranek, G. T., Reznick, J. S., Watson, L. R., & Crais, E. R. (2012). The First Year Inventory: a longitudinal follow-up of 12-month-old to 3-year-old children. *Autism*. doi: 10.1177/1362361312439633

Watson, L. R., Baranek, G. T., Crais, E. R., Reznick, S. J., Dykstra, J., & Perryman, T. (2007). The First Year Inventory: Retrospective Parent Responses to a Questionnaire Designed to Identify One-Year-Olds at Risk for Autism. *Journal of Autism and Developmental Disorders, 37*(1), 49-61. doi: 10.1007/s10803-006-0334-4

Watson, L. R., Crais, E. R., Baranek, G. T., Dykstra, J. R., & Wilson, K. P. (2013). Communicative gesture use in infants with and without autism: a retrospective home video study. *American Journal of Speech-Language Pathology, 22*(1), 25-39.

Wendland, J., Gautier, A.-C., Wolff, M., Brisson, J., & Adrien, J. L. (2010). Retrait relationnel et signes précoces d'autism: étude préliminaire à partir de films familiaux. [Relational withdrawal behaviour and early signs of autism: a preliminary study using family videos]. *Devenir, 22*(1), 51-72.

Werner, E., & Dawson, G. (2005). Validation of the phenomenon of autistic regression using home videotapes. *Arch Gen Psychiatry, 62*(8), 889-895. doi: 10.1001/archpsyc.62.8.889

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief Report: Recognition of Autism Spectrum Disorder Before One Year of Age: A Retrospective Study Based on Home Videotapes. *Journal of Autism and Developmental Disorders, 30*(2), 157-162. doi: 10.1023/A:1005463707029

Wetherby, A. M., Brosnan-Maddox, S., Peace, V., & Newton, L. (2008). Validation of the Infant—Toddler Checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism, 12*(5), 487-511. doi: 10.1177/1362361308094501

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

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

Wetherby, A., & Prizant, B. (2002). *Communication and Symbolic Behavior Scales Developmental Profile- First Normed Edition*. Baltimore, MD: Paul H. Brookes.

Wiggins, L. D., Baio, J. O. , & Rice, C. . (2006). Examination of the Time Between First Evaluation and First Autism Spectrum Diagnosis in a Population-based Sample. *Journal of Developmental & Behavioral Pediatrics, 27(2) Supplement*(2), S79-S87.

Williams, K. , Helmer, M., Mellis, C. M. , Tuck, M., Glasson, E. J. , Bower, C. I. , & Wray, J. (2005). Incidence of autism spectrum disorders in children in two Australian states. *Med J Aust, 182*(3), 108-111.

Yoder, P., Stone, W. L., Walden, T., & Malesa, E. (2009). Predicting social impairment and ASD diagnosis in younger siblings of children with autism spectrum disorder. *J.Autism Dev. Disord., 39*, 1381-1391.

Young, R. L., Brewer, N., & Pattison, C. (2003). Parental Identification of Early Behavioural Abnormalities in Children with Autistic Disorder. *Autism, 7*(2), 125-143. doi: 10.1177/1362361303007002002

Zakian, A., Malvy, J., Desombre, H., Roux, S., & Lenoir, P. (2000). [Early signs of autism and family films: a new study by informed evaluators and those unaware of the diagnosis]. *Encephale, 26*(2), 38-44.

Zappella, M., Einspieler, C., Bartl-Pokorny, K. D., Krieber, M., Coleman, M., Bolte, S., & Marschik, P. B. (2015). What do home videos tell us about early motor and socio-communicative behaviours in children with autistic features during the second year of life--An exploratory study. *Early Hum Dev, 91*(10), 569-575. doi: 10.1016/j.earlhumdev.2015.07.006

Zero to Three, National Center for Clinical Infant Programs. (1994). *Diagnostic Classification: 0-3, Diagnostic classification of mental health and developmental disorders of infancy and early childhood*. Washington, DC: Zero to Three.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23*(2–3), 143-152. doi: http://dx.doi.org/10.1016/j.ijdevneu.2004.05.001

APPENDICES

Appendix A: Study quality rating schema for the systematic review

Item	Scoring	Comment	Score
Coding scheme/ scale validated?	☐ Yes (2) ☐ No (0)		
Coders blind to DX status?	☐ Yes (2) ☐ No (0)		
IRR reported?	☐ Yes , comprehensively (2) ☐ Yes, but not clearly (1) ☐ No (0)		
Statistical analysis conducted?	☐ Yes, with effect sizes (2) ☐ Yes, no effect sizes (1) ☐ No (0)		
Controlled for video length & content?	☐ Yes, both (2) ☐ Yes, one (1) ☐ No (0)		
How was Dx status ascertained?	☐ DX assessment (2) ☐ medical records (2) ☐ screening scale (1) ☐ Parent report (1) ☐ not described (0)		
Sample size > 10/group	☐ Yes (2) ☐ No (0)		
Total quality score out of 14:			

Appendix B: Recruitment sources for study participants

B.1: Recruitment sources for parents of children with ASD

Name of source	Type of source	Method of promotion	Approximate reach
Autism Advisory and Support Services	ASD specific information and intervention service	Personal contact with parents at weekly playgroup and at special events, email to parents (twice in 12 months) , post on their Facebook page (twice in 12 months)	Direct engagement with circa 600 families
KU Autism Specific Early Learning and Care Centre kindergarten	ASD specific Kindergarten	In person by researcher	Direct engagement with 35 families
ASPECT (Autism Spectrum NSW)	ASD specific organisation	Ad on website from April 2015 until March 2016	~1000 children
ASPECT South West Sydney school	ASD specific school	email to staff and parents; staff handing out flyers to parents	~160 families
Lizard Centre	ASD specific intervention service	Flyers at reception	~120 families
ABi centre	ASD specific intervention service	Flyers at reception	~175 families
Giant Steps	ASD specific school	Ad in weekly newsletter	Direct engagement 53 families
Tony Attwood website	Asperger's resource web-site for families and clinicians	Listed on website as research project	(134 views)
Royal Far West	Service for children with special needs from non-metropolitan NSW	Clinicians handing flyers to parents	Direct engagement with 10 families
Private clinicians	3 clinical psychologists, 1 paediatrician specialising in ASD	Flyers and/or emails to parents distributed through 4 private clinicians	Direct engagement with 16 families
Australian Clinical Psychology Association	Professional organisation of clinical psychologists	Email to list serve	unknown
Total number of families reached (approximate):			**~2000 families**

B.2: Recruitment sources for general parent population

Name of source	Method of promotion	Approximate reach[1]
Essential Baby parent forum website	Ad on website	(~250,000 viewers/year)
BubHub parent forum website	Post on online parent forum	(35,000 viewers/day)
Study Facebook page	Regular posts on Facebook from October 2014	237 parents reached
KIDSize Living – online parent group	Post on Facebook page, May	(~5000 group members)
Sydney Child – print publication	Print ad in May 2015	(circulation of 127,347)
Researcher's social network	In person and via social media from October 2014 onwards	~250 people via social media profile
University of Sydney research portfolio	Listed on website, October 2015	(16,500 visitors / year)
Total number of families reached (approximate):		500+

[1] The reach for most of these sources is not specific to the study advertisement, but rather for the website or print magazine overall, thus estimations are difficult to make.

Appendix C: Survey responses - Child health and parent mental well-being questions

ID	Sex	Hearing impairment 12 m	Hearing impairment 24 m	Vision impairment 12 m	Vision impairment 24 m	Child health in first 2 years	Developmental concerns	Parental mood difficulties lasting ≥ 2 weeks in two years after child's birth Present?	Parent affected	Duration (months)	Professional help sought?
ASD1	M	No	No	No	No	Fair		Yes	both	18	Yes
ASD2	M	No	No	No	No	Very Good		No			
ASD3	M	No	No	No	No	Good		Yes	both	?	Yes
ASD4	M	No	No	No	No	Fair		Yes	father	2	Yes
ASD5	M	No	No	No	No	Very Good		No			
ASD6	M	DK	Yes	No	No	Fair		No			
ASD7	M	No	No	No	No	Good		Yes	mother	8	Yes
ASD8	M	Yes	Yes	DK	DK	Fair		Yes	mother	3	Yes
ASD9	F	No	No	No	No	Very Good		No			
ASD10	F	No	No	No	No	Fair		No			
ASD11	F	Yes	Yes	Yes	Yes	Good		Yes	mother	48	Yes
TD1	M	No	No	Yes	Yes	Very Good	Speech problems	Yes	both	1	No
TD2	F	No	No	No	No	Very Good		No			
TD3	M	No	No	No	No	Good		No			
TD4	M	No	No	No	No	Good		No			
TD5	F	Yes	No	No	No	Very Good	Speech problems	No			
TD6	M	No	No	No	No	Very Good		No			
TD7	F	No	No	No	No	Very Good		Yes	mother	5	Yes
TD8	M	No	No	No	No	Good		No			
TD9	F	No	No	No	No	Very Good		No			
TD10	M	No	No	No	No	Very Good		Yes	mother	6 -12	No
TD11	F	No	No	No	No	Very Good		No			

DK = don't know

Appendix D: Measures

D.1: The Modified Alarm Distress Baby Scale

DATE: INFANT's AGE: INFANT's NAME:

EXAMINER:

This scale is best rated by the clinician / observer on the basis of his/her observations during the clinical interview. The clinician / observer should try and socially engage the infant by smiling, chatting & touching him/her.

The rating is based on whether the infant demonstrates a given behaviour during the examination – except for eye contact and relationship, which are rated only with reference to the infant's behaviour towards the clinician / observer.

Don't rate any item if the infant spends nearly all the consult crying or is distressed.

---

1. FACIAL EXPRESSION: TOWARDS ANYONE (OR AT ANYTIME....IT DOESN'T NEED TO BE DIRECTED TO A PARTICULAR PERSON)
Assess the extent of facial expressiveness throughout the examination. Do not include crying or reactions to aversive/painful procedures (eg. oral examination) as a sign of facial expressiveness.

- □ <u>Satisfactory</u>: Facial expressiveness is clearly observed on several occasions (*3 or more*), and is either all positive (eg. smiling), or there is a range of positive and negative (eg. grimacing) expressiveness.
- □ <u>Possible problem</u>: Expressiveness is less clear, although there is a reasonable suggestion of this (positive or negative), or expressiveness is exclusively negative.
- □ <u>Definite problem</u>: There are only hints of expressiveness, expressiveness is ambiguous or absent; face appears fixed, frozen, or 'sad' for the whole period.

BUT: *If for much of the session the Infant is expressionless, yet somehow had just enough to meet the 'Satisfactory' criteria, then score it 'Possible Problem' - as clearly there is something not quite right !*

2. EYE CONTACT: TOWARDS CLINICIAN / OBSERVER ONLY
Assess the nature of eye contact towards the clinician or any other unfamiliar person. As a rough 'rule of thumb', moderate eye contact means around a second or more second; brief eye contact means less than a second; and elusive or vague means less than half a second.

- □ <u>Satisfactory</u>: At least one episode of moderate duration eye contact together with several episodes of brief eye contact.
- □ <u>Possible problem</u>: Only brief eye contact episodes, or just one moderate episode.
- □ <u>Definite problem</u>: Only one brief eye contact episode, or eye contact is vague, elusive or completely absent.

If scored as a possible or definite problem, is this behaviour different towards the parent and the clinician?
□-Yes, different □-No, not different □-Didn't assess Describe:_____
BUT: *If for much of the session the Infant 'avoids' eye contact, yet somehow had just enough to meet the 'Satisfactory' criteria, then score it 'Possible Problem' - as clearly there is something not quite right !*

3. VOCALISATIONS: TOWARDS ANYONE (OR AT ANYTIME....IT DOESN'T NEED TO BE DIRECTED TO A PARTICULAR PERSON)
Assess the amount of vocalisation, crying, and whimpering throughout the examination.

☐ Satisfactory: Several obvious, brief vocalisations, or one or two long vocalisations   (note: vocalisations may be positive or negative but do not include cries or whimpers).
☐ Possible problem: Only one or two obvious, brief vocalisations (see note above), and / or there is screaming, crying or whimpers; ambiguous vocalisations (eg., sighs or raspy sounds); or a total absence of vocalisation.

Please note that there is no 'Definite problem' response category for this item.

4. ACTIVITY: TOWARDS ANYONE (OR AT ANYTIME....IT DOESN'T NEED TO BE DIRECTED TO A PARTICULAR PERSON)
Assess head, torso, and limb movement of the infant without taking into account hands and fingers activity, both spontaneously and in response to unpleasant stimulation.

☐ Satisfactory: At least a moderate level of activity (not just in response to unpleasant stimulation).
☐ Possible problem: Reduced level of activity, or moderate level of activity occurs just in response to unpleasant stimulation.
☐ Definite problem: Very reduced level of activity regardless of the stimulation.

'Moderate' means: *reasonable arm and leg movements, on several occasions (3+), that are not just 'fairly small movements'.*

5. RELATIONSHIP: TOWARDS CLINICIAN / OBSERVER ONLY
Assess the infant's ability to engage in a relationship with the clinician or any other unfamiliar person. Relationship is assessed through the infant's eye contact and interaction with the clinician / observer.

☐ Satisfactory: Relationship at least moderately evident – either positive or negative. The infant is aware of, and displays at least a moderate social or emotional connection with the clinician / observer.
☐ Possible problem: Relationship seems tenuous or doubtful, or only seems to be evident when the infant is crying, struggling etc. There is only a minimal level of social or emotional connection with the clinician / observer.
☐ Definite problem: No relationship evident – either positive or negative.

BUT: *If for much of the session the Infant seems oblivious to the clinician/observer, yet somehow had just enough to meet the 'Satisfactory' criteria, then score it 'Possible Problem'*

If scored as a possible or definite problem, is this behaviour different towards the parent and the clinician?
☐-Yes, different   ☐-No, not different   ☐-Didn't assess   Describe:_____

CLINICIAN / OBSERVER CHARACTERISTICS
☐ Makes a good attempt to engage infant (much smiling, talking to the infant)
☐ Makes a fair attempt to engage infant (some smiling, talking to the infant)
☐ Makes a limited attempt to engage infant (little smiling, talking to the infant)

INFANT CHARACTERISTICS
☐ Infant appears to be tired ☐ Infant appears to be distressed throughout the consultation
☐ Other. *Specify*:_____

SUMMARY: # Satisfactory:_____   # Possible Problems:_____   # Definite Problems:_____

D2: The Social Responsiveness Scale – brief version (SRS-BRIEF)

SRS (BRIEF)
Please rate the truthfulness of the following statements about your child

		Not TRUE	Some-times TRUE	Often TRUE	Almost Always TRUE
1	Interprets things too literally and doesn't "get" the real meaning	0	1	2	3
2	Clumsy, not well coordinated	0	1	2	3
3	Avoids eye contact, or has unusual eye contact	0	1	2	3
4	Child has difficulty making friends, even when trying his/her best to do so	0	1	2	3
5	Has more difficulty than other children with changes in routine	0	1	2	3
6	Obsessive, thinks or talks about the same thing over and over	0	1	2	3
7	Regarded by other children as odd or weird	0	1	2	3
8	Socially awkward, even when he/she is trying to be polite	0	1	2	3
9	Has trouble keeping pace with the flow of a normal conversation	0	1	2	3
10	Difficulty "relating" to peers	0	1	2	3
11	Confuses cause and effect in a way that is inappropriate for his/her age	0	1	2	3
12	Exhibits repetitive odd behaviours such as hand flapping or rocking	0	1	2	3
13	Has difficulty answering questions directly; unintentionally talks around the subject or communicates ideas unrelated to the question	0	1	2	3
14	Walks in between two people who are talking	0	1	2	3
15	Inappropriately concentrates on parts of things rather than "seeing the whole picture"	0	1	2	3
16	Stares inappropriately; does not direct eye gaze toward the appropriate focus of attention	0	1	2	3

D.3: Online Questionnaire

Version 4 – 24 Feb 2015

Part 1 – *Screening questions (To be completed prior to up-load of videos)*

S.1 Are you a parent of a child aged 3 years or older?

☐ yes   ☐ no   -> exclude-> message will appear: *"Thank you for your interest in participating in our research study. Because we are looking for a specific group of participants you won't be able to participate in this study."*

S.2 Do you have video recordings of one of your children when he/she was around 12 months old?

     ☐ Yes, I have video recordings of my child for both time points

     ☐ Yes, I have video recordings of my child when he/she was around 12 months old

     ☐ Yes, I have video recordings of my child when he/she was around 24 months old

     ☐ No, I don't have video recordings for my child for either of those time points -> *if no, exclude [display message as above]*

S.4 Has one of your children been diagnosed with Autism or Autism Spectrum Disorder or a similar developmental disorder?

     ☐ yes   ☐ no   -> *eligible to participate in control group – skip to Participant's details ; then Part 2 Version B*

S.5 If yes, what kind of diagnosis was given to your child?

     ☐ Autistic disorder/ Autism -> eligible to participate in case group– skip to Part 1; for Part 2 *Questionnaire A*

     ☐ Asperger's Syndrome-> eligible to participate in case group– skip to Part 1; for Part 2 Questionnaire A

     ☐ Pervasive Developmental Disorder Not Otherwise Specified (PDDNOS) -> -> eligible to participate in case group– skip to Part 1; for Part 2 Questionnaire

     ☐ Childhood Disintegrative Disorder -> exclude [display message as above]

     ☐ Global Developmental delay -> exclude [display message as above]

     ☐ Developmental delay in language, motor or other areas -> exclude [display message as above]

     ☐ Other:_____ -> exclude [display message as above]

     ➔ *SRS –BRIEF questionnaire inserted here*

Participant's personal details:

     1. What is your age? _____

2. Are you ☐ Female   ☐ Male ?

3. What is your ethnicity? _____

4. What is your highest level of education:

       ☐ Year 11 or below

       ☐ HSC / completed year 12 or equivalent

       ☐ TAFE

       ☐ University

5. The research team would like to contact you in case we have any questions about the videos or anything else. To do this please let us know your preferred way of contacting you:

☐ Email address:_____   or

☐ Phone number:_____

Thank you for completing this short survey.

As we described at the start we are interested in video recordings you have made of your child, or one of your children, when they were about 12 and about 24 months old.

We will send an email with log-in details for the secure video up-load facility to you shortly; this will allow you to share videos of your child with the research team. This will also include a checklist for you to help you select a suitable video and instructions for how to up-load the videos.

A second brief survey will follow after you have up-loaded the videos.

If you have any questions at this point please contact Bettina Christl by email:
bchr2063@uni.sydney.edu.au

If you have any concerns about your child's development or require more information about autism please contact your GP, Aspect (ph: 1800 277 328), Autism Advisory and Support Services (ph 1300 222 777) or other autism association.

If participating in this study is causing you any distress please contact following support services:

- For immediate 24/7 support call Lifeline Australia on 13 11 14, or the Beyond Blue helpline on 1300 22 4636

- For information and advice about mental health and treatment talk to your GP or call the SANE Australia Helpline on 1800 18 7263 Mon – Fri, 9am – 5pm.

- For information about local support and counselling for carers of children with a disability visit the Carers Australia webpage
http://www.carersaustralia.com.au/publications/useful-links/.

Part 2 (To be completed after video up-load)

Version A – for ASD group

Thank you for having answered our questions so far and for providing [] video recordings of your child to us. We now would like to ask you some questions about the videos you provided.

1. Video 1 (your child at about 12 months)

1.1 Please provide a brief description of what your child is wearing in the video at about 12 months:_____

1.2 Are there other people (adults or children) in this video?   ☐ yes   ☐ no

1.2.1 If so, were they reasonably well known to your child at the time of the recording?

    ☐ yes, all people were reasonably well known to my child

    ☐ yes, my child knew some of the people reasonably well

    ☐ no, my child did not know any of these people very well

1.3 Can you briefly describe whether the video shows:

    ☐ a daily family routine

    ☐ a play situation

    ☐ a family gathering / special occasion

    ☐ your child's birthday

    ☐ Other; please describe: _____

1.4 Was there a reason why you selected this particular video of your child? Please describe:

_____

_____

1.5 What is the approximate date of the video recording you have chosen?

_____/_____ (MM/YYYY; e.g. 06/2010)

2. Video 2 (your child at about 24 months):

2.1 Please provide a brief description of what your child is wearing in the video at about 24 months:_____

2.2 Are there other people (adults or children) in this video?   ☐ yes   ☐ no

2.2.1 If so, are they reasonably well known to your child?

    ☐ yes, all people were reasonably well known to my child

☐ yes, my child knew some of the people reasonably well

☐ no, my child did not know any of these people very well

2.3 Can you briefly describe whether the video shows:

☐ a daily family routine

☐ a play situation

☐ a family gathering / special occasion

☐ your child's birthday

☐ Other; please describe: _____

2.4 Was there a reason why you selected this particular video of your child? Please describe:

_____

_____

__

2.5 What is the approximate date of the video recording you have chosen?

_____/_____ (MM/YYYY; eg 06/2010)

Now we would like to learn a bit more about your child:

4.   What is the date of birth of your child shown in the video recordings?
___/___/____ (DD/MM/YYYY)

If you have more than one child please enter the date of the child who has since been found to have Autism Spectrum Disorder or similar developmental problems.

5.   Is your child ☐ Female or ☐ Male?

6.   Who first seriously noticed that something may be developmentally wrong with your child?
☐   You or your partner;
☐   A relative;
☐   A health professional;
☐ Someone else. Please describe:_____

6.1 What age was your child at this time ?   _____years_____months

7.   When were 'serious' investigations for a developmental disorder started by a health professional?   Month:_____ Year:_____

8. And at what age was your child diagnosed by a health professional with Autism, Aspergers, Pervasive Developmental Disorder Not Otherwise Specified PDDNOS or similar? _____years_____months

9. Who diagnosed your child?
   - ☐ Paediatrician
   - ☐ General Practitioner
   - ☐ Paediatric psychiatrist
   - ☐ Child Developmental Unit
   - ☐ Other. Please describe:_____

10. Was the health professional(s) who diagnosed your child part of a service specialised in childhood developmental disorders?   ☐ yes ☐ no

11. How certain are you that your child has Autism, Aspergers or PDDNOS or a mix of these conditions?
    - ☐ Not at all certain
    - ☐ A little bit certain
    - ☐ Moderately certain
    - ☐ Almost certain
    - ☐ Completely certain

12. How severely impacted is your child's day to day functioning by the Autism/Aspergers/ PDDNOS?
    - ☐ No impact
    - ☐ Mild impact
    - ☐ Moderate impact
    - ☐ Severe impact

13. Did you notice any signs of developmental problems in your child before his/ her diagnosis of autism/Aspergers/PDDNOS?   ☐ yes ☐ no – > skip to 14

   13.1 If yes, how old was your child when you first noticed signs? ____years_____ months

   13.2 If yes, what signs did you notice?

    - ☐ My child was not making eye contact.
    - ☐ My child did not smile at me or anyone else.
    - ☐ My child did not respond to his/ her name.
    - ☐ My child did not point at things he/ she wanted.
    - ☐ My child started speaking very late or not at all.
    - ☐ Other signs:_____

14. Did your child seem to develop normally first and then start to lose some skills he / she had already ?   ☐ yes ☐ no -> skip to 15

   14.1 If yes, how old was your child when you noticed him / her losing some skills?
      ____years __ months

   14.2 What skills did your child lose? Please describe:_____

15. Would you have liked to have known that there was a problem with your child's development earlier than when specialist investigations begun??
      ☐ yes   ☐ no

16. Has your child experienced any other developmental problems: ☐ yes ☐ no

      16.1 If yes, please describe:_____

17. Did your child have any problems with:

      his / her hearing at around age 12 months ? ☐ yes ☐ no   ☐ Don't know

      his / her hearing at around 24 months?   ☐ yes ☐ no   ☐ Don't know

      his/ her vision at around age 12 months?   ☐ yes ☐ no   ☐ Don't know

      his/ her vision at around 24 months?   ☐ yes ☐ no   ☐ Don't know

18. How would you rate your child's overall health during his/ her first two years of live?
      ☐ Very good
      ☐ Good
      ☐ Fair
      ☐ Poor

19. Was there a period of 2 weeks or more since the birth of your child when you or your partner had mood difficulties or felt depressed?   ☐ yes   ☐ no -> skip to 20.

      a. If yes, who was having difficulties:
            ☐ I had mood difficulties
            ☐ my partner had mood difficulties
            ☐ Both my partner and I had mood difficulties

      b. If yes, when did you or your partner experience these difficulties? ____ (year; e.g 2006)

      c. How long did these difficulties last? _____ (if more than one episode, estimate total time of all episodes together; e.g 3 months)

d. Did you or your partner seek professional help for these difficulties; e.g. from your GP, a psychiatrist or psychologist? □ yes □ no

20. Would you like to receive written feedback about the observations the research team made based on the video recordings of you provided? □ yes □ no

Version B - for control group

Thank you for having answered our questions so far and for providing video recordings of your child to us. We now would like to know some more about the videos you have provided.

1. Video 1 (your child at about 12 months)

1.1 Please provide a brief description of what your child is wearing in the video at about 12 months:_____

1.2 Are there other people (adults or children) in this video? □ yes □ no

1.2.1 If so, are they reasonably well known to your child?

□ yes, all people were reasonably well known to my child

□ yes, my child knew some of the people reasonably well

□ no, my child did not know any of these people very well

1.3 Can you briefly describe whether the video shows:

□ a daily family routine

□ a play situation

□ a family gathering / special occasion

□ a child's birthday

□ Other; please describe: _____

1.4 Was there a reason why you selected this particular video of your child? Please describe: _____ _____

1.5 What is the approximate date of the video recording you have chosen?

___ /_____ (MM/YYYY; e.g. 06/2010)

2. Video 2 (your child at about 24 months):

2.1 Please provide a brief description of what your child is wearing in the video at about 24 months:_____

2.2 Are there other people (adults or children) in this video?　　☐ yes　　　☐ no

2.2.1 If so, are they reasonably well known to your child?

　　　☐ yes, all people were reasonably well known to my child

　　　☐ yes, my child knew some of the people reasonably well, but others not

　　　☐ no, my child did not know any of these people very well

2.3 Can you briefly describe whether the video shows:

　　☐ Daily family routine

　　☐ Play situation

　　☐ Family gathering / special occasion

　　☐ Child's birthday

　　☐ Other; please describe: _____

2.4 Was there a reason why you selected this particular video of your child? Please describe: _____

_____

2.5 What is the approximate date of the video recording you have chosen?

　　　　　　　　　　　　　　__/_____ (MM/YYYY; eg 06/2010)

Now we would like to learn a bit more about your child:

4. What is the date of birth of your child shown in the video recordings? __/__/____ (DD/MM/YYYY) *If more than one of your children are shown in the video please state the birth date of the child who you have been referring to throughout this survey.*

5. Is your child ☐ Female ☐ Male ?

6. Has your child experienced any developmental problems: ☐ yes ☐ no

　　6.1 If yes, please describe:_____

7. How certain are you that your child's development is normal?
　　　☐　Not at all certain

☐ A little bit certain
☐ Moderately certain
☐ Almost certain
☐ Completely certain

7.1 If moderately, a little bit or not at all certain, please describe what kind of concerns you have about your child's development:

_____

8. Did your child have any problems with

his / her hearing at around age 12 months ?	☐ yes	☐ no	☐ don't know
his / her hearing at around 24 months?	☐ yes	☐ no	☐ don't know
his/ her vision at around age 12 months?	☐ yes	☐ no	☐ don't know
his/ her vision at around age 24 months?	☐ yes	☐ no	☐ don't know

9. How would you rate your child's overall health during his/ her first two years of life?
☐ Very good
☐ Good
☐ Fair
☐ Poor

10. Was there a period since the birth of your child when you or your partner had mood difficulties such as feeling depressed or very anxious?

☐ yes   ☐ no -> *skip to question 11.*

a. If yes, who was having difficulties:

☐ I had mood difficulties
☐ My partner had mood difficulties
☐ Both, my partner and I had mood difficulties

b. If yes, when did you or your partner experience these difficulties? _____ (year; eg. 2006)

c. How long did these difficulties last? _____ (if more than one episode, estimate total time of all episodes together; e.g. 3 months)

d. Did you or your partner seek professional help for these difficulties; e.g. from your GP, a psychiatrist or psychologist?   ☐ yes   ☐ no

11. Would you like to receive feedback about the observations the research team made based on the video recordings you provided?   ☐ yes   ☐ no

End of survey message to be displayed on completion of part 2:

Thank you for taking the time to participate in our study about social interaction and communication in toddlers.

Please let me know if you have further questions in regards to this study or if you have any concerns. Feel free to contact me on bchr2063@uni.sydney.edu.au.

If you indicated that you were interested in receiving individual feedback on your child's social interaction and communication behaviour observed in the video recordings, you will receive a detailed report from us via email.

Thanks again for your participation!!

If you have any concerns about your child's development or require more information about autism please contact your GP, ASPECT (ph: 1800 277 328), Autism Advisory and Support Services (ph 1300 222 777) or other autism associations.

If participating in this study has been causing you any distress please contact following support services:

- For immediate 24/7 support call Lifeline Australia on 13 11 14, or the Beyond Blue helpline on 1300 22 4636

- For information and advice about mental health and treatment talk to your GP or call the SANE Australia Helpline on 1800 187 263, Mon – Fri, 9am – 5pm.

- For information about local support and counselling for carers of children with a disability visit the Carers Australia webpage http://www.carersaustralia.com.au/publications/useful-links/.

Appendix E: Online Participant Information Statement & Consent

Social Interaction in Toddlers (SIT) - Study

(1) **What is the study about?**

You are invited to participate in a study of social interaction and communication in toddlers with or without developmental difficulties. The study aims to investigate early social interaction and communication behaviours in infants and toddlers using home-video recordings. The study findings will help to examine how these behaviours can be measured, which we hope will improve the clinical work of early childhood nurses and paediatricians.

(2) **Who is carrying out the study?**

The study is being conducted by Bettina Christl under the supervision of A/Prof Stephen Matthey and A/Prof Cathy McMahon, and will form the basis for the degree of Doctorate/ Master of Clinical Psychology at The University of Sydney.

(3) **What does the study involve?**

Your involvement includes completing two brief online surveys about the socio-demographic status of your family and some information about your child's development.

You will be asked to provide the research team with video-recordings of your child which date from when he/she was around age 12 months and also 24 months of age.

Research participation does not involve any travel and is entirely conducted online.

If you wish to receive individual feedback about your child's behaviour observed in the video-recording then this will be given to you.

Participation in this study poses no identifiable risk to you or your child.

(4) **How much time will the study take?**

Each of the two questionnaires will take about 10 minutes to complete. Up-loading the two video-recordings to the secure file-sharing facility (CloudStore) will take about 30 minutes.

(5) **Can I withdraw from the study?**

Being in this study is completely voluntary - you are not under any obligation to consent and - if you do consent - you can withdraw at any time without affecting your relationship with The University of Sydney, the Autism Advisory and Support Services (AASS) or Aspect. Your responses and your video recordings will then be removed from the study and deleted.

(6) **Will anyone else know the results?**

121

All aspects of the study, including results, will be strictly confidential and only the researcher team will have access to information on participants except as required by law. A report of the study results may be submitted for publication, but individual participants will not be identifiable in such a report.

(7) Will the study benefit me?
You may gain insight into your child's early social interaction and communication behaviour. However, we cannot and do not guarantee or promise that you will receive any benefits from the study.

(8) Can I tell other people about the study?
Yes, you can tell other people about the study and direct them to either the study's Facebook page or the researcher's contact details if they wish to find out more.

(9) What if I require further information about the study or my involvement in it?
When you have read this information, you can contact Bettina Christl with any questions you may have now or at any other stage: bchr2063@uni.sydney.edu.au.

(10) What if I have a complaint or any concerns?
Any person with concerns or complaints about the conduct of a research study can contact The Manager, Human Ethics Administration, University of Sydney on +61 2 8627 8176 (Telephone); +61 2 8627 8177 (Facsimile) or ro.humanethics@sydney.edu.au (Email).

☐ **I give consent to my participation in the research project Social Interaction in Toddlers (SIT) Study. In giving my consent I acknowledge that:**

- The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.

- I have read the Participant Information Statement and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s.

- I understand that being in this study is completely voluntary – I am not under any obligation to consent.

- I understand that my involvement is strictly confidential. I understand that any research data gathered from the results of the study may be published however no information about me will be used in any way that is identifiable.

- I understand that I can withdraw from the study at any time, without affecting my relationship with the researcher(s) or the University of Sydney or the Autism Advisory and Support Services (AASS) now or in the future.

☐ **I do not give consent to my participation in this research project.**

If you have any concerns about your child's development or require more information about autism please contact your GP, ASPECT (ph: 1800 277 328), Autism Advisory and Support Services (ph 1300 222 777) or other autism associations.

If participating in this study is causing you any distress please contact following support services:

- For immediate 24/7 support call Lifeline Australia on 13 11 14, or the Beyond Blue helpline on 1300 22 4636

- For information and advice about mental health and treatment talk to your GP or call the SANE Australia Helpline on 1800 187 263, Mon – Fri, 9am – 5pm.

- For information about local support and counselling for carers of children with a disability visit the Carers Australia webpage http://www.carersaustralia.com.au/publications/useful-links/.

Appendix F: Ethics Approvals

F.1: Approval of Original submission

Research Integrity
Human Research Ethics Committee

Wednesday, 23 April 2014

Dr Caroline Hunt
Psychology; Faculty of Science
Email: caroline.hunt@sydney.edu.au

Dear Caroline

I am pleased to inform you that the University of Sydney Human Research Ethics Committee (HREC) has approved your project entitled "**Social Interaction in Toddlers (SIT study) - Examining the clinical utility of the modified Alarm Distress Baby Scale (m-ADBB) for the detection of early signs of Autism Spectrum Disorder.**".

Details of the approval are as follows:

Project No.:   2014/124

Approval Date: 23 April 2014

First Annual Report Due: 23 April 2015

Authorised Personnel: Hunt Caroline; Christl Bettina; Matthey Stephen; McMahon Cathy;

Documents Approved:

Date Uploaded	Type	Document Name
29/01/2014	Advertisements/Flyer	Study advertisement for online parent forums
29/01/2014	Other Instruments/Tools	EA scales 4th edition
28/01/2014	Other Instruments/Tools	m-ADBB scale
6/04/2014	Other Type	Video instruction email
27/01/2014	Participant Consent Form	SIT Study_Online Participant Info Statement and Consent form
6/04/2014	Questionnaires/Surveys	Online Questionnaire
27/01/2014	Recruitment Letter/Email	SIT study_Participant invitation email template

HREC approval is valid for four (4) years from the approval date stated in this letter and is granted pending the following conditions being met:

Condition/s of Approval

- Continuing compliance with the National Statement on Ethical Conduct in Research Involving Humans.

- Provision of an annual report on this research to the Human Research Ethics Committee from the approval date and at the completion of the study. Failure to submit reports will result in withdrawal of ethics approval for the project.

- All serious and unexpected adverse events should be reported to the HREC within 72 hours.

F.2: Approval of amendment #1

THE UNIVERSITY OF SYDNEY

Research Integrity
Human Research Ethics Committee

Thursday, 31 July 2014

Dr Stephen Matthey
Psychology; Faculty of Science
Email: stephen.matthey@sswahs.nsw.gov.au

Dear Stephen

Your request to modify the above project submitted on 4 July 2014 was considered by the Executive of the Human Research Ethics Committee at its meeting on **23 July 2014**

The Committee had no ethical objections to the modification/s and has approved the project to proceed.

Details of the approval are as follows:

Project No.:　　　　**2014/124**

Project Title:　　　　**Social Interaction in Toddlers (SIT study) - Examining the clinical utility of the modified Alarm Distress Baby Scale (m-ADBB) for the detection of early signs of Autism Spectrum Disorder.**

Approved Documents:

Date Uploaded	Type	Document Name
04/07/2014	Questionnaires/Surveys	Participant Questionnaire - revised-clean
04/07/2014	Questionnaires/Surveys	SRS-brief - questionnaire

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

Dr Stephen Assinder
Chair
Human Research Ethics Committee

F.3: Approval of amendment #2

THE UNIVERSITY OF SYDNEY

Research Integrity
Human Research Ethics Committee

Wednesday, 4 March 2015

Dr Stephen Matthey
Psychology; Faculty of Science
Email: stephen.matthey@sswahs.nsw.gov.au

Dear Stephen

Your request to modify the above project submitted on 15 January 2015 was considered by the Executive of the Human Research Ethics Committee at its meeting on 17 February 2015.

The Committee had no ethical objections to the modification/s and has approved the project to proceed.

Details of the approval are as follows:

Project No.:   2014/124

Project Title: Social Interaction in Toddlers (SIT study) - Examining the clinical utility of the modified Alarm Distress Baby Scale (m-ADBB) for the detection of early signs of Autism Spectrum Disorder.

Approved Documents:

Date Uploaded	Type	Document Name
24/02/2015	Questionnaires/Surveys	Online Questionnaire - revised clean
24/02/2015	Participant Info Statement	Participant information statement revised - clean

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

Dr Fiona Gill
Chair
Executive, Human Research Ethics Committee

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2007), NHMRC and Universities Australia Australian Code for the Responsible Conduct of Research (2007) and the CPMP/ICH Note for Guidance on Good Clinical Practice.

F.4: Approval of amendment #3

THE UNIVERSITY OF SYDNEY

Research Integrity
Human Research Ethics Committee

Monday, 19 October 2015

Dr Stephen Matthey
Psychology; Faculty of Science
Email: stephen.matthey@sswahs.nsw.gov.au

Dear Stephen

Your request to modify the above project submitted on 10 July 2015 was considered by the Executive of the Human Research Ethics Committee at its meeting on 21 July 2015.

The Committee had no ethical objections to the modification/s and has approved the project to proceed.

Details of the approval are as follows:

Project No.:   2014/124

Project Title: Social Interaction in Toddlers (SIT study) - Examining the clinical utility of the modified Alarm Distress Baby Scale (m-ADBB) for the detection of early signs of Autism Spectrum Disorder.

Approved Documents:

Date	Type	Document
14/10/2015	Advertisements/Flyer	Recruitment email - final appeal
08/09/2015	Recruitment Letter/Email	Clinician Letter/Email to parents
08/09/2015	Advertisements/Flyer	Study Flyer Version 3
10/07/2015	Participant Consent Form	Clinician Consent Form
10/07/2015	Participant Info Statement	Clinician Information Statement
06/07/2015	Advertisements/Flyer	SIT Study Flyer version 3 clean
06/07/2015	Other Type	Invitation letter or email

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

Dr Lilon Bandler
Chair
Human Research Ethics Committee

F.5: Ethics approval from ASPECT

Bettina Christl

From:	Kerry Ritzrow <kritzrow@autismspectrum.org.au>
Sent:	Wednesday, 6 May 2015 8:27 AM
To:	'bchr2063@uni.sydney.edu.au'
Subject:	Aspect Research Approval for amendments

Good morning Bettina,

The research committee has looked over your amendments for your project *Examining the clinical utility of the modified Alarm Distress Baby Scale (m-ADBB) for the detection of early signs of Autism Spectrum Disorder* and are pleased to approve the study becoming a Tier 2 application with approval.
Please contact me about getting in touch with the principals of the schools. We have 8 schools and it would be great for you to let me know which ones you would like to contact and how you would like to contact them. I can contact them on your behalf or I can write a letter indicating that you will be approaching the schools individually. The approval of the schools for you to advertise your study is entirely at the principals discretion.
Please let me know how you would like to proceed.

Thanks
Kind regards
Kerry

Go Colourful for World Autism Awareness month this April #coloursforautism

Go Colourful for World Autism Awareness Month this April! #coloursforautism

Kerry Ritzrow Aspect Practice Research & Administration Assistant
4/3 Aldgate St & Blacktown Rd
Prospect NSW 2148
T 02 8868 8504 **M**
www.autismspectrum.org.au

Aspect acknowledges Australia's traditional owners. We pay our respects to them and their elders both past and present.

Disclaimer: If you are not an authorised recipient of this email, please contact Autism Spectrum Australia (Aspect) immediately by return email or by telephone on 61 2 8977 8300. In the case that you are not an authorised recipient, you should not read, print, transmit, store or act in reliance on this email or any attachments, and should destroy all copies of them. This email and any attachments are confidential and may contain legally privileged information and/or copyright material of Autism Spectrum Australia (Aspect) or third parties. You should only transmit, distribute or commercialise the material if you are authorised to do so.

1