

Early Detection of Autism Spectrum Symptomatology in Very Preterm Infants

Charlotte Sanderson

June 2016

Submitted in partial fulfilment of the requirements for the degree of Doctor in
Clinical Psychology (DClinPsy), Royal Holloway University of London.

Acknowledgements

Thanks to my supervisors, Neil Marlow, Jane Vosper and Michelle de Haan, for your support and guidance throughout this process. I am grateful to have been able to learn from your expertise and experience.

Thank you to the PDP team at UCL, most notably Kayleigh Day and Jade Okell, for your help in collecting this data. Thank you for helping me to meet my seemingly impossible deadlines, and for making this endeavour a less lonely process!

Thank you, Mum and Dad. I can't put into words the help you have been over the past few months. Thanks for being at the end of the phone whenever I needed you, and for looking out for me always. But don't ever let me do any more degrees!

Thanks to my amazing friends, Jeff and Slim, for listening, for supporting and for your "motivational limericks" when I needed them the most.

Thank you to my fellow trainees. I will look back so fondly on these last three years because of you all. Thank you for keeping the laughter going even through the never-ending weekends in the library.

Thank you Anthony, for your patience, understanding and encouragement, throughout this process and always.

Most importantly, sincerest thanks to the families and children who have made this research possible.

Abstract

Children born very preterm (VP) are at an increased risk for a wide range of neurodevelopmental impairments, including autism spectrum disorder (ASD) and related symptomatology. The present thesis investigated early behavioural markers for ASD in VP infants. Using a prospective, longitudinal design, three early infancy markers which have been linked to ASD outcomes in other high-risk groups were investigated: the Autism Observation Scale for Infants (Bryson, Zwaigenbaum, Mcdermott, Rombough, & Brian, 2008), a disengagement of visual attention task (Mayada Elsabbagh, Fernandes, et al., 2013) and infant temperament profiles (Putnam, Helbig, Gartstein, Rothbart, & Leerkes, 2014). Existing data from 65 full term and 41 VP infants at 6 and 12 months was collated, audited and examined using a series of bivariate and multivariate analyses. A sub-set of these infants were followed up at 30-34 months to estimate associations with ASD outcomes, as indicated by the Autism Diagnostic Observation Schedule (ADOS II) (Lord et al., 2012).

VP infants showed greater impairment on the AOSI, a behavioural assessment of multiple ASD markers, compared to full-term controls at both 6 and 12 months. These impairments appeared to be independent of developmental delays. AOSI scores at 12 months also showed association with ASD outcomes at 30-34 months. VP infants were also rated by their parents as showing more *negative affect* during the first year of life, though no association with ASD outcomes was observed. These findings are discussed in relation to the wider literature on early detection in ASD. In particular, the likelihood of heterogeneity in early trajectories towards ASD symptoms is discussed. Multiple marker assessments like the AOSI may be particularly useful tools for detecting ASD symptomatology earlier in high-risk infant populations, in order to enable earlier targeted intervention. Future research may consider the longer-term stability of these early profiles in VP infants, and explore specific biological and psychosocial risk factors for emerging ASD symptomatology in this group.

List of Tables

Table 1 Overview of the longitudinal study protocol, including existing data previously collected at 6 & 12 months, and new follow-up data collected at 30-34 months.	49
Table 2 Description of the 18 behavioural ‘markers’ coded in the AOSI	57
Table 3 Total Number of Infants participating in the study at 6, 12 and 30-34 months with N’s for Individual Measures.....	75
Table 4 Overall sample characteristic of the VP and full-term control groups, including all infants contributing data at 6m, 12m and/or 30-34 months .	77
Table 5 Descriptive Statistics for AOSI Total Scores and Total Markers at 6 and 12 months for VP and full-term infants.....	79
Table 6 Composite Scores for Cognitive, Language and Motor Subscales on the Bayley III	85
Table 7 Model A: Simple linear regression including birth status only	87
Table 8 Model B: Hierarchical linear model including Bayley cognitive scores (Step 1) and birth status (Step II)	88
Table 9 Model C: Hierarchical linear model including Bayley Motor scores (Step 1) and birth status (Step II)	88
Table 10 AOSI positive screening rates (7+ markers) at 6 and 12 months....	89
Table 11 Attention Disengagement Summary Scores at 6 and 12 months....	91
Table 12 IBQ-R-VSF Total and Subscale Scores at 6 and 12 months	93
Table 13 Sample characteristics for VP and full-term infants who completed the ADOS II at 30-34 months	97
Table 14 ADOS Total and Subscale Scores at 30-34 months	98
Table 15 Correlation matrix summarizing Pearson R correlations between key early infancy measures and ADOS II (30-34m) and Bayley III (24m) outcomes.....	102

List of Figures

Figure 1 Proportion of live births per year in the UK according to completed weeks of gestation (Centre for National Statistics; 2011)	33
Figure 2 Stimuli (screenshots) of attention disengagement task showing the three trial types: Baseline (A), Gap (B) and Overlap (C)	61
Figure 3 Graph showing significant main effects of group and time-point for AOSI scores	81
Figure 4 Illustration of marginally significant difference ($p=.022$) in parent reported Negative Affect at 12 months of age.....	95

Table of Contents

Chapter 1: Introduction	9
Overview of Autism Spectrum Disorder	10
Early Intervention in ASD	14
Detecting ASD Earlier: Research into ‘markers’ during Infancy	16
Atypical Early Social Skills.....	17
Atypical Early Attentional Skills	21
Atypical Early Temperament	25
Multiple Marker Clinical Observation Tools: The Autism Observation Schedule for Infants (AOSI).....	28
Summary and Identification of Gaps in the Early Detection Literature	30
Overview of Preterm Birth	33
Associations between preterm birth and ASD	35
Why are preterm infants at risk for ASD? An integrative model	37
Neurobiological Factors: Neonatal Brain Abnormalities	37
Psychosocial and Environmental Factors.....	39
Preterm Birth as a “Double Hazard”	40
Summary	41
Early Detection of ASD in VP Infants	41
Rationale and Outline of the Study	43
Study Hypotheses	46
Chapter 2: Method	47
Ethics	47
Design	48
Power Analyses and Sample Size Estimations	50
Sample Characteristics	52
Very Preterm (VP) Group	52
Inclusion/Exclusion Criteria.	52
Recruitment.	53
Sample Overview.	53
Correction of Age for Prematurity.	53
Full Term Controls	54
Inclusion/Exclusion Criteria.	54
Recruitment.	54
Sample Overview.	54
Measures	55
The Autism Observation Scale for Infants (AOSI)	55
Overview.....	55
Attention Disengagement Eye-tracking Task	59
Infant Behaviour Questionnaire Revised Very Short Form (IBQ-R-VSF)	64
Autism Diagnostic Observation Schedule – Second Edition (ADOS-II)	66
Bayley Scales of Infant and Toddler Development - Third Edition (Bayley III) ..	68
General Procedure	69

Collection and Audit of Existing Infancy Data.....	69
Collection of New Data (ADOS II) at 30-34 months.....	70
Data Analysis Procedure	71
Data Exploration	71
Cross-Sectional Comparisons	71
Longitudinal Associations	72
Service User Involvement.....	72
Chapter 3: Results.....	74
Sample Characteristics.....	74
Demographic Matching.....	78
Cross Sectional Comparisons	78
Autism Observation Scale for Infants	79
Bivariate Analysis of AOSI Continuous Scores.....	80
Estimating the Influence of Gender and SES	82
Hierarchical Regression	84
Comparison of AOSI Positive Screening Rates	89
Attention Disengagement	90
Data Exploration.....	90
Temperament (IBQ-R-VSF).....	93
Estimating the Influence of Possible Confounds: Gender and SES.....	95
Autism Diagnostic Observation Schedule (ADOS II).....	96
Bivariate Comparison of Continuous Scores.....	97
Comparison of Diagnostic Cut-offs.....	99
Exploration of Longitudinal Associations Relating to ASD	99
Selection of Appropriate Statistical Methods.....	99
Exploratory Partial Correlation.....	103
Summary of Findings from Correlational Analyses	104
Chapter 4: Discussion	105
Overview of Findings Relating to Study Hypotheses	105
Autism Observation Scale for Infants (AOSI)	107
Cross Sectional Comparisons	107
Associations between AOSI Scores and ADOS II Outcomes	109
Summary of Findings Relating to the AOSI.....	113
Attention Disengagement	113
Summary of Findings Relating to Attention Disengagement.....	116
Infant Behaviour Questionnaire (Temperament)	116
Conceptual and Methodological Challenges Associated with ‘Infant Temperament’	120
Summary of Findings Relating to Infant Temperament.....	122
ASD Symptomology at 30-34 months (ADOS II).....	122
ASD Diagnostic Categorisations	123
DSM V: Discriminating Social Affect and Repetitive Behaviours.....	126
Summary of Findings on the ADOS III	128
Strengths and Limitations	129
Strengths of the Study	129

Limitations of the study	131
Sample Sizes.....	131
Group Matching (SES & Gender).....	134
ASD Outcome Measurement.....	136
Theoretical and Clinical Implications	137
Recommendations for Further Research.....	140
Dissemination	143
Concluding Statements	143
References	145
Appendices	161
Appendix 1 – NHS Ethics Approval Letter	161
Appendix 2 – NHS Ethics Substantial Amendment Approval Letter (A)	164
Appendix 3 – NHS Ethics Substantial Amendment Approval Letter (B)	166
Appendix 4 – Royal Holloway Ethics Approval (Email)	168
Appendix 5 – Table of A-priori Power Calculations	169
Appendix 6 – Letter to Nurseries (30-34 month recruitment).....	170
Appendix 8 – Autism Observation Schedule for Infants (AOSI)	176
Appendix 9 – Infant Behaviour Questionnaire Revised – Very Short Form (IBQ-R-VSF).....	177
Appendix 10 – Autism Diagnostic Observation Schedule II.....	178
Appendix 11 - Longitudinal Study Parental Consent Form.....	179
Appendix 12 – 30-34 Month Only Parental Consent Form (full term)	180
Appendix 13 – Parent Study Experience Feedback Form	181
Appendix 14 – Demographic Questionnaire Section for Parent Education Level	182

Chapter 1: Introduction

Autism spectrum disorder (ASD) is a relatively common neurodevelopmental disorder that has pervasive impacts upon social communication skills, emotions and behaviour. Though it is still rarely diagnosed before the third year of life (Yirmiya & Charman, 2010), there is evidence that earlier targeted intervention may lead to better functional, behavioural and emotional outcomes (Bryson, Rogers, & Fombonne, 2003; Warren et al., 2011). Understanding the earlier signs of ASD is therefore a priority to enable more timely detection and intervention.

Prospective research with infants at high familial risk for ASD has now highlighted various social and non-social 'markers' visible within the first year of life, which may represent the earliest signs of the disorder (Jones, Gliga, Bedford, Charman, & Johnson, 2014; Yirmiya & Charman, 2010). However, it is not clear whether these early behavioural markers generalise to other high-risk infant groups. Infants born very preterm (VP) are also known to be at elevated risk for developing ASD (S. Johnson & Marlow, 2011; Kuzniewicz et al., 2014; Treyvaud et al., 2013), as well as wide-ranging neurodevelopmental impairments (Hutchinson et al., 2013). Extending the prospective study design to VP infants therefore offers an opportunity to test whether these same behavioural markers are associated with ASD *specific* outcomes in more globally at-risk groups. The present study therefore seeks to investigate early behavioural markers for ASD in VP infants, using a prospective, longitudinal design.

This chapter will first provide an overview of the relevant literatures surrounding ASD and VP birth. First, key concepts underlying ASD will be outlined, including an overview of diagnosis, epidemiology, comorbidity and long-term outcomes. The literature surrounding early diagnosis and detection of ASD in the wider population will be critically reviewed, with a particular focus upon findings from prospective, longitudinal research with high risk 'infant sibling' populations. The following section will then provide an overview of VP birth, particularly focusing upon the apparent links between prematurity and ASD. The relatively limited literature surrounding early detection of ASD in preterm populations will be discussed, highlighting the particular challenges of early screening for ASD in globally 'at-risk' infants. These two literatures will then be drawn together, to propose how a prospective approach with VP infants could further our understanding of early detection of ASD in preterm groups, and in the wider population. Specific study aims and study hypotheses will be presented.

Overview of Autism Spectrum Disorder

Autism Spectrum Disorders (ASD) are neurodevelopmental disorders characterised by difficulties with social interaction and communication, as well as restrictive and repetitive patterns of interest and/or behaviour (DSM-5) (American Psychiatric Association, 2013). Though estimates of prevalence vary, it is thought to affect approximately 1% of the general population (Baird et al., 2006). The risk of reoccurrence within families is, however, much higher

than this, with reoccurrence rates of 10% to 19% in those families with at least one child already diagnosed with ASD (Carter & Scherer, 2013).

Although genetic linkages have been found (Rutter & Thapar, 2014), the etiology of ASD is still not fully understood. Twin and family studies suggest that genetic factors play a substantial role, with concordance rates of around 36% amongst monozygotic twins, compared to around 3% in dizygotic twins (Lauritsen & Ewald, 2001). However, a monozygotic concordance rate of well below 100% also implicates significant non-genetic influences. Corroborating this, a range of pre- and peri-natal risk factors have been linked to ASD (Larsson et al., 2005) including birth complications (such as breech presentation and low Apgar score at five minutes), parental psychiatric history and, most importantly for this research, preterm birth.

Until the introduction of DSM-V (American Psychiatric Association, 2013), 'ASD' was an umbrella term used to represent a range of related diagnoses, including autistic disorder, pervasive developmental disorder—not otherwise specified, and Asperger's disorder. However, DSM-V now recognises just one single diagnostic category of 'autism spectrum disorder' in order to capture the significant variability in presentation that is known to exist between individuals with the condition. DSM-V also now specifies that in order to reach diagnostic criteria, an individual must present with *both* difficulties in social functioning (i.e. social-emotional reciprocity, nonverbal communicative behaviours and developing and maintaining relationships), *and* restricted and repetitive patterns of behaviour. These 'non-social' features may include

stereotyped or repetitive speech, movement or use of objects (e.g. spinning or flicking objects; rocking), excessive adherence to routines/rituals or resistance to change, restricted or rigid interests, and hypo- or hyper-sensitivity to environmental stimuli.

Diagnosis of ASD is still made principally on the basis of clinical presentation (Yirmiya & Charman, 2010) with assessment typically constituting structured behavioural observation and a detailed clinical interview with caregivers. At present, diagnosis is most often made between two and three years of age (Charman & Baird, 2002). First concerns are however often raised by parents before 18 months, and it is increasingly acknowledged that some features – such as lack of social smiling, limited eye contact and absence of appropriate facial expressions - can be evident as early as 8-12 months (Osterling & Dawson, 1994). This can leave parents with a challenging and frustrating delay between initial concerns about their child's development, and receiving a thorough assessment and appropriate support.

Children with ASD are also more likely than their peers to have *other* additional needs or neurodevelopmental impairments, including difficulties with attention, mood, cognitive functioning and adaptive skills (Charman & Baird, 2002; Hanson et al., 2013; Perry, Flanagan, Geier, & Freeman, 2009). Mannion and Leader (2013) highlight that the nature of comorbidity changes across the lifespan. Young infants with ASD are more likely than their typically developing peers to present with feeding and sleeping difficulties (Kozlowski, Matson, Belva, & Rieske, 2012), and higher levels of avoidance and anxiety

(Davis et al., 2010). At school age, children with ASD are more likely to meet diagnostic criteria for social anxiety, attention-deficit/hyperactivity disorder and oppositional defiant disorder (Simonoff et al., 2008). Into adulthood, clinically significant difficulties with attention, anxiety and mood continue to be more prevalent in these individuals (Mannion & Leader, 2013).

The social and economic implications of ASD across the lifespan are significant, with a lifetime 'cost' of support per individual estimated at approximately £1.5 million (Buescher, Cidav, Knapp, & Mandell, 2014). A key factor in this is the need for special education provisions and parental productivity loss during childhood. As many as 85% of individuals with ASD continue to have cognitive and/or adaptive difficulties that prevent them from living independently into their adult lives, necessitating the use of residential services and/or extensive continued support from parents and families (Howlin, Goode, Hutton, & Rutter, 2004). The burden of having a child with ASD is thus often pervasive and lifelong for parents, particularly in terms of navigating complex and ever-changing services (Pottie & Ingram, 2008).

In summary, ASD represents a spectrum of social and non-social difficulties that have pervasive, wide-ranging and lifelong impacts. Though it affects approximately 1% of the wider population, certain populations are at dramatically increased risk for developing these difficulties.

Early Intervention in ASD

Accumulating evidence now underlines the potential for early intervention to be beneficial for children with ASD (Estes et al., 2015; Granpeesheh, Dixon, Tarbox, Kaplan, & Wilke, 2009; Rogers et al., 2012; Zachor, Ben-Itzhak, Rabinovich, & Lahat, 2007). Although the most effective form of early intervention remains unclear (Camarata, 2014), numerous studies of parent-led interventions have demonstrated marked improvements in *targeted* skills including social communication and imitation skills and global gains in adaptive and cognitive functioning (Zwaigenbaum, Bryson, & Garon, 2013) and reduced stress for families (Kaaresen, Rønning, Ulvund, & Dahl, 2006). Earlier intervention also appears to be more cost and time efficient than a “wait and see” approach (Chasson, Harris, & Neely, 2007). Koegel, Koegel, Ashbaugh, and Bradshaw (2014) argue that this is largely linked to the prevention of emerging secondary symptoms, such as challenging behaviours (e.g. tantrums/self-injury) or mood difficulties, which are often directly related to difficulties with communication and/or socialisation.

Particularly promising outcomes have been reported from the Early Start Denver Model (ESDM) - a comprehensive developmental behavioral programme for 18-30 month olds (Waddington, van der Meer, & Sigafos, 2016). The ESDM broadly aims to enhance social and affective engagement using a three-part programme, including provision of training to parents or therapists, intensive one-to-one intervention, *and* group intervention. Multiple studies have now reported positive child, parent, and therapist outcomes

using this model (Waddington et al.). Most notably, a randomised controlled trial (Dawson et al., 2010) reported that infants enrolled on the programme for two years showed significant improvements in IQ, language, adaptive behavior and ASD symptomology relative to children receiving routine community intervention. Moreover, later follow up of these children indicated that benefits typically appeared to be sustained after the termination of therapeutic input. At age six, it was reported that levels of ASD symptomology and challenging behavior had actually continued to improve since the end of therapy. This suggests that interventions like ESDM have the potential to have a marked and lasting impact upon development. However, it is notable that this kind of intervention is highly intensive, and without further research, may lack viability in typical community settings.

Despite the potential benefits of earlier intervention, it is still very rare for ASD to be diagnosed before the age of two (Yirmiya & Charman, 2010).

Developmental models of neurodevelopmental disorders and brain plasticity would indicate that the earlier a targeted intervention is commenced, the greater the potential for altering atypical trajectories (Karmiloff-Smith, 1998). It is argued that if we can detect early difficulties or markers before the full 'syndrome' associated with ASD has emerged, it may be possible to develop interventions that limit the impact of early perturbations to neurodevelopment. Dawson and colleagues (2015) argue that this is because very early intervention has the potential to modify "*risk processes*", such as altered social interactions, that are triggered by early environmental or genetic

susceptibility, and which progressively lead to adverse neurodevelopmental outcomes or 'syndromes'. Together these findings highlight that efforts to promote earlier detection in ASD are a priority.

Detecting ASD Earlier: Research into 'markers' during Infancy

Research into the 'prodrome' of ASD (i.e. signs and symptoms that are visible before the full 'syndrome' emerges) has been dominated by retrospective and prospective study designs (Yirmiya & Charman, 2010). Retrospective reports from parents with children with ASD, and analysis of behaviour in old home videos has indicated various patterns of behavior visible within the first year of life that appear to predict later ASD development (Barbaro & Dissanayake, 2009). For example, retrospective analysis of 'first birthday party' videos has suggested that limitations in joint attention, gaze avoidance and social gaze may be earlier markers for later diagnosis (Werner, Dawson, Munson, & Osterling, 2005). Retrospective studies, however, are fundamentally flawed in their lack of standardised measurement; data is broadly restricted to that which is memorable to parents and/or is caught on film, and as such, systematic biases are likely, including an increased likelihood of positive behaviours.

Prospective study designs which longitudinally follow the development of 'at-risk' infants offer a more systematic and controlled approach to investigating early markers for ASD (Jones et al., 2014). A number of groups have to date carried out large-scale longitudinal, prospective studies of this type, with infants at familial risk for ASD (Jones et al., 2014; Zwaigenbaum et al., 2013).

If a child has an older sibling diagnosed with ASD, there is approximately a 1-in-5 chance that they will also reach criteria by their third birthday (Ozonoff et al., 2011)¹. By following development in these high-risk siblings and in low-risk controls from early infancy until an age where a reliable diagnosis can be made (2-3 years), associations between early markers and ASD outcomes can be explored. These prospective studies point towards a range of behavioural markers observable during infancy, that appear to be associated with emerging ASD symptomatology in high-risk infant siblings (Jones et al.). The following section will selectively review findings from this extensive infant sibling literature, focusing upon three domains that have received significant attention: early social skills, early attentional skills and early “temperament”.

Atypical Early Social Skills

Qualitative impairments in social interaction are a core diagnostic feature of ASD, and deficits are typically observed in both initiation and response to social interaction. The key diagnostic features relate to atypical use of nonverbal behaviors to regulate social interaction (e.g. mutual gaze, facial expression, posture and gestures); failure to develop peer relationships; lack of spontaneous seeking to share enjoyment or interest (e.g. showing, giving

¹ The likely reasoning for this being higher than predicted by community samples (c. 10%, Constantino et al., 2010) is a combination of “stoppage effects” (parents choosing not to have additional children when one child has a disability) and milder forms of ASD not being detected in the community.

and pointing); and lack of social or emotional reciprocity (e.g. not participating in simple social play).

In typical development, infants start to demonstrate skills in initiating and responding to social interaction within months of birth (Trevvarthen & Aitken, 2001). Social smiling normally starts to emerge as early as 1-2 months (Anisfeld, 1982), and infants typically start developing simple facial imitations, like tongue protrusions, from birth. Imitation continues to be a powerful tool for social and cognitive learning, and infants start to imitate simple actions with objects from around 6 months (Jones et al., 2014). Over the first year, infants also become increasingly sensitive to gaze cues from others, showing reliable following of gaze from around 10 months (Jones et al., 2014). Also by around 9 months, a baby will start to use gestures (pointing/showing) and gaze to initiate periods of joint attention, and show interest in simple social games (e.g. peekaboo) (M. H. Johnson et al., 2005).

Prospective studies have indicated that high-risk infant-siblings later diagnosed with ASD show typical rates of social smiling, and just as much interest in looking at their mothers faces as typically developing infants during the first year of life (Young, 2009). One key social-communicative initiation skill that *does* appear to be affected, however, is gesture. Landa and colleagues (2007) found that 14 months olds who went on to receive a diagnosis of ASD by 30-36 months have a smaller repertoire of gestures, and show fewer initiations of joint attention than those who do not. Prospective and retrospective parent-report also highlights reduced gesture use as a

predictor of later diagnosis (Osterling & Dawson, 1994; Zwaigenbaum et al., 2005). Jones et al. (2014) argue that early gestural deficits might have a cascading effect upon cognitive and social communication development in ASD because it is a key tool for eliciting information from social partners.

Studies have also indicated that high-risk infants who go on to receive an ASD diagnosis are less socially responsive in early infancy. For example, Macari et al. (2012) completed the toddler module of the Autism Diagnostic Observation Schedule (ADOS-T) (Luyster et al., 2009) with high-risk infant siblings and low-risk controls at 12 months, and found lower levels of engagement with the researcher in high-risk infants who were later diagnosed. Similarly, Sullivan et al. (2007) found that consistent failure to respond to gaze and point cues at 14-months was predictive of 36 month ASD diagnosis. Bedford and colleagues (2012) agreed that atypicalities in response to eye-gaze may be an important marker. They found that 13 month-old infants who went on to develop ASD did not look longer at items others' referenced with their gaze. These findings suggest that difficulties understanding the referential importance of eye-gaze as a source of social information might be an early marker for ASD.

Evidence from prospective studies also suggests that atypical imitation may be an important early marker for ASD outcomes. For example, Zwaigenbaum et al. (2005) found that difficulties with object-action imitation (such as copying an adult tapping an object) on the semi-structured 'Autism Observation Scale for Infants' (AOSI) (Bryson, 2008) assessment at 12 months predicted

diagnosis of autism at 24 months. In a recent study looking prospectively at parent reported social communication skills, Rowberry et al. (2015) also found that imitation impairments reported at 12 months were linked to ASD outcomes at 24 or 36 months, with large effect sizes. Notably, in this study, parent-reported *decline* in play and communication alongside impaired vocal imitation predicted later developing ASD with high specificity. However, the specificity of atypical imitation as a marker for emerging ASD, as opposed to other adverse outcomes, has been questioned. In a comprehensive prospective study of imitation skills (including actions on objects, facial gestures and hand gestures), infant siblings who went on to be diagnosed with ASD showed equivalent imitation impairments to those who went on to have *other* developmental difficulties; *both* of these groups of infants showed imitation difficulties relative to low-risk controls (Young et al., 2011). This highlights potential problem of specificity in delineating early trajectories of ASD – in order to be helpful, markers need to be both sensitive and specific to ASD.

Another indication of the importance of early social processing skills, and particularly the processing of eye-gaze, in emerging ASD comes from research using event related potentials (ERP). Elsabbagh and colleagues (2012) measured neurophysiological responses to eye-gaze shifts in 6-10 month old infants and found atypicalities in the 'P400' response (an early brain response to visual stimuli) were associated with later emerging ASD. Specifically, they found that infants who met diagnostic criteria for ASD at 36

months did *not* show enhanced P400 responses when an adult shifted their gaze away from them, whereas typically developing infants did. These neurophysiological studies underline the importance of considering atypical early social processes in the early detection of ASD, and indicate that subtle atypicalities in brain responses may even be detectable before atypicalities in social behaviour are observable. However, it is notable that the specificity of these ERP findings to the processing of *social* information is uncertain.

Atypical Early Attentional Skills

Other research with high-risk ASD siblings has highlighted the role of atypical visual attentional processes in emerging ASD (Sacrey, Armstrong, Bryson, & Zwaigenbaum, 2014). Visual attention refers to the ability to direct and sustain visual focus on an action or event, and is the primary means of both exploring the world and self-regulation during infancy (M. H. Johnson, 1990). The process of visual orienting (moving attention from one object in the environment to another) is understood to comprise three cognitively and neuronally separate phases: disengagement, shifting and engagement of attention. Each of these phases has been shown to map onto specific and different brain regions, with ‘disengagement’ mediated by the parietal cortex, ‘shift’ associated with the midbrain, and ‘engagement’ linked with the thalamus (Posner, Walker, Friedrich, & Rafal, 1987; Posner, 2014; Rees, 2009). The visual orienting network is known to undergo dramatic development through the first year of life (Hood, 1995); “sticky fixation” or “obligatory looking” is a typical phenomenon in young infants, and is thought

to represent a difficulty with the top-down processes of attention disengagement (Hood & Atkinson, 1993). Reaction times for visual disengagement from one of two competing stimuli decreasing significantly between two and six months of age (M. H. Johnson, Posner, & Rothbart, 1991; McConnell & Bryson, 2005) and then continues to improve or refine through to approximately ten years (Wainwright & Bryson, 2005).

Visual orienting is a critical skill for infants, as it allows them to explore their environment and regulate levels of arousal by disengaging from sources of over or under stimulation (M. H. Johnson et al., 1991). As a result, it is theorised that disruptions to processes of disengagement may have significant and wide-ranging impacts on development (Johnson et al., 2005). The impacts of early impairments in disengagement have been discussed in relation to various domains, such as novelty processing, arousal, joint attention, social attention, and executive functioning (Griffith, Pennington, Wehner, & Rogers, 1999; Johnson et al., 2005). It is argued that such impairments may *disproportionately* affect social development, because the attentional demands of social information processing are so high (Griffith, Pennington, Wehner, & Rogers, 1999; M. H. Johnson et al., 2005). As a result, attention disengagement skills have been a focus in ASD research, and particularly in the search for early 'markers' for emerging ASD.

There is converging evidence to suggest that young *children* with ASD do present with impairments in disengaging visual attention, and may actually show similar 'sticky fixation' to much younger infants (Landry & Bryson, 2004).

Studies have typically measured disengagement using a computer-based task (often referred to as the “Gap Overlap” or “Gap” task), which measures the latency to begin an eye movement or ‘saccade’ from a central stimulus, after the appearance of a peripheral stimulus. The task involves two types of trials: ‘baseline’ trials, where the central stimulus simply disappears when the peripheral stimulus appears, and experimental ‘overlap’ trials. On overlap trials the central stimulus remains on-screen when the peripheral target appears, so shifting attention involves effortful disengagement. The difference between performance on baseline and overlap trials can then be calculated to estimate ‘disengagement effect’.

Various prospective studies have now completed this “gap-overlap” task with high-risk infants. Two of these (Elsabbagh et al., 2013; Zwaigenbaum et al., 2005) found that though the task was not discriminatory of ASD outcome at 6 months, infants diagnosed with ASD at 24-36 months actually got *slower* at disengaging attention between 6 and 12 months. In contrast, high-risk typically developing infants and low-risk infants tended to get faster at disengaging. It appeared to be this deterioration in visual disengagement skills in the second half of the first year that predicted later ASD diagnosis. Elison et al. (2013) also reported evidence that early deficits in disengagement of attention may be linked to emerging ASD. They reported slower latencies to shift attention on overlap trials as early as 7 months in infants who later scored ‘high’ on the ADOS at 25 months, compared to both high-risk typically developing infant-siblings and low-risk infants. A number of

authors have therefore argued that early impairments in disengagement of attention could be the *primary* impairment in a developmental model of ASD (e.g. Elsabbagh et al., 2009; Landry & Bryson, 2004). Models have suggested that early difficulties flexibly disengaging visual attention affect the development of other functions like arousal regulation and joint attention downstream, which in turn contribute to the emergence of ASD symptoms.

There remain a number of limitations to these arguments. First, if disengagement of attention was the primary early deficit in ASD, then these difficulties should be observable *before* any other behavioral markers appear – but there is little evidence that they do (Jones et al., 2014). For instance, Bedford and colleagues (2014) found that difficulties with disengagement and joint attention (gaze following) at 14 months made independent and additive contributions to diagnostic outcome at 36 months, rather than disengagement predicting joint attention or vice versa. That is, although attentional disengagement skills were predictive of ASD outcomes, it was those children that also independently showed difficulties with joint attention that were most likely to develop ASD. This may suggest that tools that measure multiple domains, or patterns of deficits, will likely be the most effective approach to earlier identification – as they capture the additive effects of multiple impairments. The second major limitation with early disengagement of attention theories of ASD is that the evidence lacks specificity at present; it is unclear from prospective research with ASD infant-siblings alone whether these early impairments are associated with ASD specifically, or whether they

are a generic marker for 'atypical' development (Jones et al., 2014; Roncadin et al., 2011). Research with other high-risk infant groups, who are at risk for a wider range of adverse developmental outcomes (including ASD), could shed light on this question, and test the generalisability of attention disengagement impairments as a marker for ASD.

Atypical Early Temperament

The concept of 'temperament' received considerable attention in the infant development literature towards the end of the last century, and has more recently gathered interest as a possible early infancy marker for ASD (del Rosario, Gillespie-Lynch, Johnson, Sigman, & Hutman, 2014). Though there is some disagreement over exactly what infant temperament measures, it is typically conceived as a set or profile of stable behavioral dimensions affecting attention, behavioral reactivity, emotion regulation and activity level (Thomas & Chess, 1977; Zwaigenbaum et al., 2005). The conceptual overlap between dimensions of 'temperament' and some of the behaviours linked to the autistic phenotype (such as high emotional reactivity and poor adaptation to novelty or change) has led researchers to consider it as a potentially useful construct for measuring early signs or markers for ASD in high-risk infants.

Measurement of early temperament in high-risk infant populations has typically relied on parent report measures, such as the Infant Behaviour Questionnaire (Gartstein & Rothbart, 2003) or the Early Childhood Behavior Questionnaire (Putnam, Gartstein, & Rothbart, 2006). These questionnaires tend to summarise temperament profiles according to a number of different

dimensions, including negative affect, positive affect (or “surgency”) and effortful regulation of emotion and/or attention. Research using these scales has suggested that a range of atypicalities in early temperament may be visible when comparing ASD infant siblings to full-term infants (Clifford et al., 2013; Garon et al., 2009) – and that these profiles may be *most* atypical in those infant siblings who actually go on to develop ASD (Zwaigenbaum et al., 2007). There appears, however, to be relatively poor agreement between studies in terms of what specific temperament profile best ‘predicts’ ASD outcomes.

Using the IBQ-R, Clifford and colleagues (2013) observed reduced surgency (i.e. positive affect) and poorer self-regulatory skills in high-risk ASD siblings compared to controls during the first two years of life. However, it was a profile of increased perceptual sensitivity, increased negative affect and reduced tolerance of physical affection (‘cuddliness’) that predicted the high-risk infants who reached diagnostic criteria for ASD at 36 months.

Zwaigenbaum and colleagues (2005) also observed temperament abnormalities in a high-risk sibling group, but found a slightly different pattern of impairment predicted ASD outcomes. They found that the infants who later reached diagnostic criteria for autism showed marked passivity and decreased activity levels at 6 months, and extreme distress reactions and decreased expression of positive affect by 12 months. A somewhat different pattern again was reported by Garon and colleagues (2009) using the Child Behaviour Questionnaire (CBQ) at 24 months. They found that high-risk infant

siblings were rated as having more difficulty with “effortful emotion regulation” (higher negative affect, lower positive affect, and difficulty controlling attention and behavior) and fewer ‘approach behaviours’ (i.e. responsiveness to “social” reward cues) relative to low-risk control infants – though only the latter predicted actual ASD outcomes amongst the high-risk group.

Thus although there do appear to be meaningful relationships between early temperament and ASD outcomes, the precise nature of these profiles remains somewhat unclear. Del Rosario and colleagues (2014) argue that this may be because ASD is linked to distinctive temperament trajectories during infancy, rather than stable impairments. The difficulty in finding accurate markers within the construct of temperament may also reflect the challenges of relying on parent report; altered reference points for judging their child’s temperament might be a particular source of bias in infant sibling populations (Rothbart & Mauro, 1990).

In a similar vein to measures of attention disengagement, the specificity of ‘atypical temperament’ as a potential marker for ASD remains unclear. Some evidence suggests that it may be more of a non-specific marker for atypical development. Using a parent-report measure of infant temperament, Garon et al. (2009) found that though high-risk siblings who were later diagnosed with ASD (at 36 months) showed poorer effortful emotional regulation skills than low-risk infants, they had only marginally poorer skills than infants who later developed other atypical outcomes. This highlights a key limitation in the current early ASD literature. By relying almost universally on ASD infant

sibling populations in prospective research, it is very difficult to know whether atypicalities in temperament (and other markers) observed during early infancy are *specifically* predictive of ASD, or whether they simply reflect atypical or delayed development more generally. It is also unclear whether atypical temperament would predict ASD outcomes in other high-risk groups. In terms of understanding the early trajectories of ASD and enabling earlier detection, these problems of generalisability and specificity in infant markers are significant. Further prospective research with different high-risk infant groups, and particularly those who are at risk of broader developmental impairments such as preterm infants, could help to fill these problematic gaps in the literature.

Multiple Marker Clinical Observation Tools: The Autism Observation Schedule for Infants (AOSI)

Research with high-risk infant siblings has increasingly pointed towards patterns or profiles of behaviour, rather than isolated impairments, being the most potentially useful approach to early detection in ASD (Jones et al., 2014; Bedford et al., 2014). There are various reasons why this might be the case. Bedford and colleagues highlight the role of additive effects in terms of amplifying risk, with infants who have multiple deficits in non-social and social domains appearing to be the most likely to develop clinical difficulties.

In recognition of this thinking, one research group has developed a systematic clinical observation tool to measure a range of possible markers for ASD in infants, based upon the findings from retrospective and prospective research

with infant siblings. The Autism Observational Scale for Infants (AOSI; Bryson, Zwaigenbaum, Mcdermott, Rombough, & Brian, 2008) is a semi-structured, play-based and researcher-led behavioural assessment, which is designed to measure early behavioural markers of ASD in infants between 6 and 18 months. Using series of presses, the assessment provides a context for the elicitation and observation of various ASD-related markers. Broadly, these relate to social communication (e.g. anticipatory social response, social babbling, orientation to name, eye contact), non-social behaviours (e.g. disengagement of visual attention, motor control, atypical sensory behaviours) and temperament (e.g. reactivity, ease of transitions between activities).

There are various advantages of assessment tools like the AOSI. First, as a play-based assessment, it arguably offers a more ecologically valid measure of early social and non-social skills than more experimental, computerised tasks. Secondly, as a clinician led assessment, it does not rely on parent report – permitting greater objectivity in ratings. Thirdly, and importantly, the simultaneous assessment of a series of potential markers makes it possible to capture the additive impact of multiple deficits or impairments occurring together. By considering multiple markers in combination, such approaches are arguably more able to capture and tolerate heterogeneity in profiles. As a result, the AOSI has the potential to be a particularly sensitive measure for detecting early signs of ASD across different high-risk groups.

It appears that total scores on the scale at 12 months, but not at 6 months, may be a promising predictor of ASD outcomes at 24 months (Zwaigenbaum

et al., 2005). More recently, Gammer and colleagues (2015), found that AOSI scores at 14 months (but, similarly, not at 7 months) were moderately correlated with later scores on the Autism Diagnostic Observation Schedule (ADOS) in a study with high-risk infant siblings (Lord et al., 2000). They also found that AOSI scores in high-risk infants who went on to reach diagnostic criteria were higher than both low-risk siblings and high-risk siblings who developed 'typically'. Importantly, Gammer and colleagues detected that scores from specific social *and* non-social items, such as orienting to name and visual-tracking, were also significantly higher in the ASD-group at 14 months. Such findings are important as they indicate the potential to translate the output of more experimental prospective research into validated and clinically relevant assessment tools.

It is important to note that at present, the AOSI has only been validated for use within high-risk ASD infant sibling populations. As such it is unclear how well the tool will generalise to other high-risk infants groups as a measure of ASD *specific* outcomes in more diverse infant populations.

Summary and Identification of Gaps in the Early Detection Literature

In summary, there has been a substantial amount of research into early infancy markers and trajectories associated with ASD, which aim to enable earlier detection and targeted intervention. Prospective research with the infant siblings of children with ASD has particularly highlighted the role of early atypicalities with social interactions, attention disengagement and infant

temperament. Multiple marker assessment tools, such as the AOSI, which aim to measure a range of these markers simultaneously may be particularly effective means of detecting emerging ASD.

It is likely that the early markers for ASD are not homogenous across infants (Jones et al., 2014). It is therefore of note that prospective research has almost universally recruited high-risk infant groups that comprise younger siblings of children diagnosed with ASD. Not only does this suggest that findings may be specific to infants at increased genetic risk, but they specifically relate to “multiplex families” – that is, families with multiple children affected by ASD (Jones et al., 2014). There is evidence that this group are *not* necessarily representative of the wider ASD population. It has been reported that cases of ‘regression’ are unusually rare in prospective sibling studies compared to known rates in the wider ASD population. As regression is relatively common in the wider ASD population (approximately 20-40% of infants appear to show losses of previously acquired skills; Baird et al., 2008), there is a definite risk that infant siblings are not representative of other high-risk groups, or the wider ASD population. Therefore, in order to characterise the generalisability of these early markers beyond ‘infant siblings’, it is very important to consider other high-risk groups prospectively.

Paediatric research over the last decade has suggested that another high-risk group for ASD is those born very prematurely. Indeed, it appears that very preterm infants are as much as ten times more likely than their term born peers to meet criteria for ASD (Eaton, Mortensen, Thomsen, & Frydenberg,

2001; Indredavik et al., 2004; S. Johnson et al., 2010b; Larsson et al., 2005; Schendel & Bhasin, 2008). However, this group is slightly different, as they are a group of infants at risk for a wide range of adverse developmental outcomes. High-risk preterm populations are therefore an important group to study, as they offer the opportunity to test the specificity of the discussed early infancy “ASD” markers to ASD, as opposed to other neurodevelopmental impairments.

The following section will first introduce key concepts relating to very preterm birth, and provide a brief overview of current understanding from the paediatric and epidemiological literature. This will particularly relate to developmental and neurodevelopmental outcomes after very preterm birth. The potential to expand ‘high-risk’ infant research to the study of early development and ASD in preterm populations will then be discussed.

Overview of Preterm Birth

Preterm birth is defined by the World Health Organisation (WHO) as any birth occurring before completion of 37 weeks of gestation, which accounts for approximately 7% of UK live births (ONS, 2011). The reasons for a baby being born early are varied and complex, but spontaneous and indicated (i.e. via caesarean section or induction) preterm birth typically relate to maternal, placental and/or fetal infection or disease (Ferrara et al., 1994). Over recent decades, the limits of viability (i.e. the gestational age at which babies can survive) have decreased dramatically (Luu et al., 2009), meaning that a significant number of babies are now born in the very preterm (VP; <32 weeks) or extremely preterm (EP; <28 weeks) range and survive (see Figure 1).

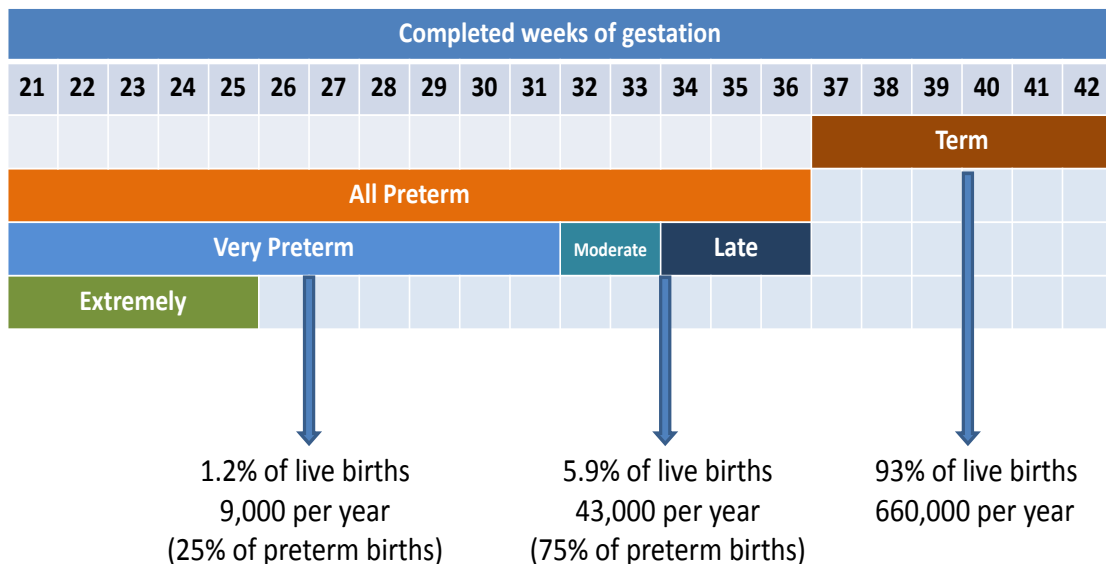


Figure 1 Proportion of live births per year in the UK according to completed weeks of gestation (Centre for National Statistics; 2011)

Although more infants are surviving VP birth, it is increasingly recognised that these babies are at an elevated risk for a wide range of adverse outcomes in the long term (Hutchinson et al., 2013; Vicari, Caravale, Carlesimo, Casadei, & Allemand, 2004). Not only are these children more likely to be diagnosed with a severe disability, including cerebral palsy or a profound cognitive or neurosensory impairment (Espy et al., 2003; Marret et al., 2013), they also appear to be at high risk for a wide range of neurodevelopmental sequelae affecting cognitive, behavioural, social and emotional development (Böhm, Smedler, & Forssberg, 2004; Foster-Cohen, Edgin, Champion, & Woodward, 2007; Saigal & Doyle, 2008). Imaging research indicates that this increased risk is linked to, in part, a vulnerability to perinatal brain injury with premature birth (S. P. Miller et al., 2005).

Though outcomes are highly heterogeneous amongst preterm samples (Day, Van Lieshout, Vaillancourt, & Schmidt, 2015), the evidence unsurprisingly indicates that the earlier a child is born, the greater the risk for brain injury (i.e. white matter damage, intraventricular hemorrhage and cortical/deep grey matter damage) and neurodevelopmental problems (Isaacs et al., 2004; Marlow, Wolke, Bracewell, & Samara, 2005; Saigal & Doyle, 2008). In the absence of cerebral injury, those born after 32 weeks have only a minimally increased risk of neurodevelopmental impairment (Gurka, LoCasale-Crouch, & Blackman, 2010) – whereas around three quarters of those born at 26-31 weeks (i.e. VP) will present with either cognitive, educational or behavioural difficulties in childhood, and more than half will have difficulties in multiple

domains (Hutchinson et al., 2013). It is estimated that the mean full scale IQ (FSIQ) of children born at 28 weeks is 11 points lower than that of their term-born class-mates, with greater deficits in children born at <26 weeks (MacKay, Smith, Dobbie, & Pell, 2010).

Research has indicated that there may be a relatively unique behavioural profile associated with prematurity, most strongly characterised by attentional difficulties and ‘internalising’ behaviours (Bhutta, Cleves, Casey, Cradock, & Anand, 2002; Johnson & Marlow, 2011). Converging evidence is also now indicative of wide-ranging impacts of very preterm birth on children’s social functioning as they grow up (S. Johnson, 2015). In a systematic review, Ritchie, Bora and Woodward (2015) outline an increased risk for social withdrawal and peer relationship problems, as well difficulties with emotional regulation and compliance in school and peer settings. Furthermore, and of particular note for this research, children born VP or EP are also considerably more likely to meet diagnostic criteria for a social communication disorder (i.e. Autism Spectrum Disorder) (Johnson & Marlow, 2011).

Associations between preterm birth and ASD

Recent research has suggested that preterm infants, as a group, are as much as ten times more likely than their term born peers to meet criteria for ASD (Eaton et al., 2001; Indredavik et al., 2004; S. Johnson et al., 2010b; Larsson et al., 2005; Schendel & Bhasin, 2008). In a British whole-population follow-up of 219 children born extremely preterm, Johnson and colleagues (S. Johnson et al., 2010b) detected rates of ASD as high as 16% at 8-11 years old. The

majority of these children met criteria for the DSM-IV diagnosis of “Autism”, which means that they presented with both social communication difficulties *and* restricted and repetitive behaviours, and as such would meet meet DSM-V criteria for ASD (Smith, Reichow, & Volkmar, 2015). There also appears to be a clear quantitative elevation in ASD symptomatology, amongst those who do not fully meet criteria. Using a parental screening questionnaire (The Child Symptom Inventory), Hack and colleagues (2004) compared 219 extremely low birthweight (ELBW) and 176 normal birth weight children at age three, and found that the ELBW group presented with more frequent, and more severe, ASD symptoms. Similarly, Johnson and colleagues (2010a) found a quantitative elevation of ASD-related symptoms compared to controls on the Social Communication Questionnaire (parent-reported) in their EP ‘EPICure’ cohort.

Aside from highlighting social communication as an area of difficulty for VP/EP survivors, this pattern of findings suggests that ASD-type symptoms amongst preterm children operate on a similar, though elevated, ‘spectrum’ to that seen in the wider population (S. Johnson et al., 2011). The association between ASD and broader neurodevelopmental impairments in high-risk preterm populations is, however, notable. In the EPICure cohort, ASD symptom severity was independently associated with both withdrawn behaviour and functional disability at 2.5 years, and with cognitive impairment, inattention/hyperactivity and peer problems by the age of 6 years (S. Johnson et al., 2010b). It is possible that VP/EP infants present with a range of

impairments that overlap, but are not identical to, those described in 'idiopathic' ASD and infants at genetic risk for the disorder. This has important implications for questions of early identification, as it might predict that the early behavioural markers for "ASD" symptomology are measurably different to those seen in infant sibling populations.

Why are preterm infants at risk for ASD? An integrative model.

Though it is now firmly established that VP/EP birth is a perinatal risk factor for ASD (amongst other neurodevelopmental difficulties), the reasons *why* this is the case appear less well understood. Models of neurodevelopment and brain plasticity propose that even subtle perturbations to brain development in early life, when neuronal connectivity is forming at a rapid rate, can lead to significantly altered developmental trajectories and emerging, specific deficits (Karmiloff-Smith, 1998). Preterm birth is associated with increased rates of neonatal brain injury, high exposure to stressors (including painful procedures) and a very abnormal early visual, sensory and social experience during a critical period of brain development. It is therefore unsurprising that infants who endure the stresses of preterm birth are at increased risk for a wide range of difficulties. The evidence for some of these factors being associated with ASD-specific risks are briefly explored.

Neurobiological Factors: Neonatal Brain Abnormalities

The relationship between neonatal brain injury, aberrant cerebral development and later neurodevelopmental impairments in VP/EP children is complex and far from fully understood. Some key findings are, however,

highlighted. Magnetic resonance imaging (MRI) at term equivalent age has consistently shown that diffuse white matter injury is common amongst preterm infants, and that the presence of such injury is correlated with poorer developmental outcomes (Dyet et al., 2006); white matter abnormalities have in fact been shown to be a better predictor of later neurodevelopmental impairment than gestational age (e.g. Woodward, Anderson, Austin, Howard, & Inder, 2006), particularly in relation to motor and cognitive outcomes (Iwata et al., 2012; Woodward et al., 2006). These findings suggest that structural and functional brain abnormalities may offer some of the earliest indicators of poor prognosis in high-risk preterm populations.

Few studies have examined the link between neonatal brain alterations, including white matter abnormality, and later ASD symptomology specifically in VP populations. One research group reported that cerebellar hemorrhage and reduced cerebellar volume on neonatal MRI were associated with positive screens on the Modified Checklist for Autism in Toddlers (M-CHAT) (Limperopoulos et al., 2007; Limperopoulos, Robertson, Sullivan, Bassan, & du Plessis, 2009). More recently, Ure and colleagues followed up 172 VP children at age seven, and found associations between ASD diagnosis and cystic lesions in the cortical white matter and reduced cerebellar volume at birth. Localised structural abnormalities have also been noted in the orbitofrontal cortex and anterior cingulate on the neonatal MRIs of EP children who go on to be diagnosed with ASD (Padilla et al., 2015). In the general literature, these brain regions have been consistently linked to executive and

behavioural regulation deficits and to ASD more specifically (Bechara, Damasio, & Damasio, 2000; Girgis et al., 2007; Jones et al., 2014; Russell, 1997). Together, these findings suggest that structural brain abnormalities likely play a role in the increased risk for ASD in preterm populations.

Psychosocial and Environmental Factors

Other authors have highlighted complex relationships between atypical environmental and psycho-social factors and preterm brain development (Perlman, 2001, 2003). VP/EP birth results in lengthy hospitalisation of infants who are physiologically unprepared for stressors outside the protective intrauterine environment. The first few weeks of life for a preterm infant on the neonatal intensive care unit (NICU) involve extended exposure to light and noise, acute and chronic illness, maternal separation, invasive procedures, handling, and multiple medications and pain-inducing procedures. Preliminary evidence suggests that such stressors during the preterm period may result in long-term alterations to more generalised stress systems (e.g. basal cortisol levels) (Grunau et al., 2007). The impacts of such changes on long-term neurodevelopment are not yet well understood, but recent research has linked NICU stress with regional alterations in brain structure and function (G. C. Smith et al., 2011). In particular, abnormal sensory exposure in the NICU environment and atypical early mother-infant contact is linked to adverse neurobehavioral and cognitive outcomes (Pineda et al., 2014).

Preterm Birth as a “Double Hazard”

Epidemiological research has shown that premature births occur more frequently in disadvantaged social environments, characterised by lower levels of parental education, low maternal age and high unemployment rates (Nadeau, Tessier, Boivin, Lefebvre, & Robaey, 2003; L. K. Smith, Draper, Manktelow, Dorling, & Field, 2007). This has led several authors to describe preterm infants as a “double hazard” population, with neurobiological and environmental risk having an additive effect upon risk for social and behavioural difficulties (Escalona, 1982; Parker, Greer, & Zuckerman, 1988). Sameroff & Chandler (Sameroff, Chandler, & Horowitz, 1975) historically reported that outcomes in low-birthweight children are closely tied to the general socio-economic circumstances into which they were born, which they referred to as the ‘continuum of caretaking casualty’. It is not clear, however, how well this research translates into modern day preterm populations, and particularly those infants now born at the lowest end of viability (VP/EP).

It is also of note that no clear association is established between social economic status and ASD outcomes either. Though some studies argue that low social economic status (SES) *is* associated with higher risk for ASD in the general population (Rai et al., 2012), other epidemiological studies have reported either no association with SES (Larsson et al., 2005) or even *decreased* prevalence of ASD in low SES groups (Durkin et al., 2010). It has been argued that this latter finding may be more of a reflection of unequal access to services, particularly in areas of the USA, rather than differential

risk per se (Rai et al., 2012). There is no clear evidence, therefore, to state that increased rates of ASD in preterm populations are linked to the socio-demographic features of this infant group.

Summary

The relatively limited research available suggests that the specific risk-factors for ASD in VP populations is likely to be complex, multifaceted and unique to this high-risk group. It is likely that there are complex inter-relationships between biological and psychosocial risk factors that place VP/EP infants at particularly high risk for atypical social affective development and ASD type symptomatology. There is a clear need for further research into these complex relationships.

Early Detection of ASD in VP Infants

VP children are a group who typically have complex needs, and as such early intervention for specific difficulties, such as ASD, is arguably particularly important. After leaving the neonatal unit, NICE recommends that VP infants are carefully monitored over the first few years of life using standardised developmental assessments (NICE, 2015). However, early screening for ASD specifically has so far proved very problematic in this group (Limperopoulos et al., 2008), and no routine screening of social communication development is currently recommended in child development services following VP/EP birth. A key reason for this is the apparent failure of available screening tools to discriminate early ASD symptomatology from other features of global

impairment, which are common in this group (T. Moore, Johnson, Hennessy, & Marlow, 2012).

Various studies have attempted to assess the utility of the Modified Checklist for Autism in Toddlers (M-CHAT) (Pinto-Martin & Levy, 2004; Robins, Fein, Barton, & Green, 2001) and similar parent-report screening tools in two year olds who were born VP or earlier (Kuban et al., 2009; Limperopoulos et al., 2008; O'Shea et al., 2009). These studies initially suggested that as many as 25% present with autistic features in toddlerhood – a number that significantly exceeds the positive screening rate in the general population (~5%). Major neurodevelopmental impairments (motor, cognitive, visual, and hearing) were, however, shown to account for as many as 50% of positive M-CHAT screens detected. This raises clear concerns over the discriminant validity and specificity of early screening tools for identifying ASD traits in VP/EP infants (Yirmiya & Charman, 2010).

It is possible that clinician-led infant assessments, such as the Autism Observation Scale for Infants (AOSI) would be more resilient to neurodevelopmental impairments than parent screening questionnaires. To our knowledge, only one study has considered the performance of a preterm group on the AOSI. In a conference submitted paper, Roncadin and colleagues (Roncadin et al., 2011) reported that infants born late to moderately preterm (i.e. <37 weeks) scored higher on the AOSI than low risk, full-term controls, but *lower* than high-risk infant siblings. Preterm infants' AOSI scores were also moderately correlated with later ASD symptomatology

at 24 months. These findings suggest that the AOSI may be more effective than basic parent questionnaire approaches at isolating ASD specific difficulties from wider developmental delays in preterm groups. However, a key limitation of this study was the relatively *low-risk* preterm population investigated. As discussed, risk for ASD in individuals born late-moderately preterm, whilst still raised, is considerably lower than in VP/EP populations (De Jong, Verhoeven, & van Baar, 2012; Kuzniewicz et al., 2014), and these infant are less likely to have wider neurodevelopmental impairments (Goldenberg, Culhane, Iams, & Romero, 2008). As such it is unclear whether the AOSI would be a useful tool for measuring specific ASD-related difficulties in higher-risk preterm samples.

Rationale and Outline of the Study

The prospective study of infants born VP could help to fill several problematic gaps in the current literature relating to early ASD development. Firstly, as VP/EP infants represent another high-risk group for developing ASD symptomology, the population provides a test for the generalisability of specific markers and multiple-marker infant assessments like the AOSI beyond ASD infant-siblings. It is reasonable to expect that the trajectories towards ASD (and thus the early markers for the disorder) are not homogeneous across different infant populations with very different initial risk factors (i.e. perinatal vs. genetic). Thus, markers that show associations with ASD in infant-sibling groups, may not in VP infants. Secondly, due to the wider developmental risks associated with VP/EP birth, the prospective study

of these infants can also challenge the specificity of markers to ASD outcomes relative to more generalised developmental delay. In turn, better understanding of the early markers for ASD-type difficulties in VP/EP groups has the potential to inform earlier targeted intervention in this vulnerable group.

To address these questions, this study therefore proposes to extend a high-risk infant prospective research design to a VP population. The study seeks to investigate three measures that have been highlighted as potential indicators of emerging ASD in the infants-sibling literature: attentional disengagement; early infant temperament; and a play-based multi-marker assessment of early non-social and social behaviour (the AOSI) (Bryson et al., 2008). In line with research carried out with ASD infant siblings, this study will adopt a prospective, longitudinal design, using data collected at three separate time points (6, 12 and 30-34 months) from infants born very preterm (VP; ≤ 31 weeks gestation) and a group of infants born at full-term (full-term; 37-41 weeks). The study will make use of both existing data at 6 and 12 months, and newly collected follow-up data at 30-34 months.

Data from 6 and 12 month assessments with VP and full-term infants has already been collected as part of a wider existing research programme.

Though detailed audit of raw data-sets is required, these existing data are available for a sample of 65 full-term and 41 VP infants on three separate measures: an attention disengagement eye-tracking task; a measure of infant temperament the IBQ-R (Very Short Form) (Putnam et al., 2014) and the

AOSI (Bryson et al., 2008). All VP and full-term infants had also completed the Bayley Scales of Infant and Toddler Development - Third Edition (Bayley III) (Bayley, 2006) at 12 months, which provides an estimate of their level of cognitive, motor and language development. VP infants had also completed the Bayley III at 24 months as part of their routine outpatient developmental follow-up. Outputs from all of these assessments will be collated and examined to address the research questions below.

Infants who are old enough to be followed up at 30-34 months will be invited back for follow-up to complete a semi-structured clinician led Autism Diagnostic Observation Schedule assessment (ADOS-II) (Lord et al., 2012) to estimate ASD symptom presence. The ADOS II was selected as an ASD outcome measure for a number of reasons. Firstly, it is considered to be one of the “gold standards” in assessment of ASD symptomology in children and in adults (Lord et al., 2012). As a behavioural observation measure, which is administered and coded by trained clinicians, it can be considered more objective than parent report measures. Therefore if practical constraints restrict assessment to one modality, a direct behavioural assessment is often considered the most reliable approach to estimating ASD symptomology (Lord et al., 2012). Another advantage of the ADOS II compared to other measures is that it estimates symptoms directly in relation to diagnostic criteria outlined in DSM V (American Psychiatric Association, 2013). As such, it provides indicators of diagnostic cut-off, as well as separate summary scores for the two core domains of the ‘ASD’ profile as it is currently conceptualised: social

affect (SA) and restricted and repetitive behaviours (RRB). This will permit some consideration of the specific profile of social and non-social difficulties presented by VP infants.

Study Hypotheses

Using the outlined design, the study seeks to address the following research hypotheses:

- 1) Compared to infants born at full-term, VP infants will show atypicalities on three 'early marker' measures at 6 and 12 months: the Autism Observation Schedule for Infants (AOSI), an attention disengagement task and parent-reported infant temperament.
- 2) Children born VP will show elevated scores relative to full term controls on the ADOS II (a standardised behavioural assessment of ASD symptomology) at 30-34 months.
- 3) It is expected that total scores on the AOSI at 12 months, but not at 6 months, will be associated with ASD symptomatology at 30-34 months in the VP group. A similar pattern of association with outcomes is tentatively predicted on measures of early attention disengagement and infant temperament.

In addition to the hypotheses, the role of developmental delays (motor; language and cognitive) in group differences and longitudinal associations in the VP group will be explored.

Chapter 2: Method

Ethics

This study is part of a wider research programme at a specialist London hospital investigating development after VP birth. This single-site programme includes brain imaging during the perinatal period and longitudinal developmental follow-up of infants born VP and term born controls. Ethical approval for the wider research programme, including the study addressed in this project, was granted by the North West London Research Ethics Committee (REC reference number 10/H0720/80). A further two amendments were also subsequently approved, the first permitting follow-up between two and three years of age and the second approving the recruitment of an additional group of term-born control participants who would take part at this 2-3 year follow-up only. Ethical approval for this study was also granted by the Department of Psychology at Royal Holloway University of London. Corresponding ethics documentation is shown in Appendices 1-4.

Parents of infants taking part were offered no payment for their participation, though reasonable expenses were provided for food/travel associated with visits. Various strategies were used to aid retention of infants between longitudinal follow-ups including newsletters, a Facebook page and a small gift for participating infants (t-shirts or teddy bear).

Design

The study comprised a prospective, longitudinal design comparing two groups (VP infants and full-term controls) at three time-points: 6m, 12m and 30-34m.

This study design was informed by prospective research being carried out by the British Autism Study of Infant Siblings (BASIS) (e.g. Elsabbagh et al., 2013), and used existing data already collected at the 6 and 12 month time-points as well as new outcome data collected at 30-34 months. The complete longitudinal study protocol is shown in Table 1.

Table 1 *Overview of the longitudinal study protocol, including existing data previously collected at 6 & 12 months, and new follow-up data collected at 30-34 months.*

Measure	Description	6m	12m	Follow-Up
Autism Observation Scale for Infants (Bryson et al., 2008)	Semi-structured, play based behavioural assessment of multiple potential early markers for ASD	x	x	
Attention Disengagement Task (Elsabbagh et al., 2009)	Eye-tracking task measuring difficulties disengaging visual attention	x	x	
Infant Behaviour Questionnaire – Revised (Very Short Form) (Putnam, Helbig, Gartstein, Rothbart, & Leerkes, 2014)	37 item parent report measure of infant temperament. Yields three summary scores: Negative Affect; Surgency (Positive Affect); Effortful Control (Self-regulation).	x	x	
Autism Diagnostic Observation Schedule (ADOS II) (Lord et al., 2012)	Semi-structured play-based diagnostic assessment of Autism Spectrum symptomatology			30-34 months
Bayley Scales of Infant and Toddler Development - Third Edition (Bayley, 2006)	Structured, play-based assessment of infant cognitive, motor and language development		x	24 months (<i>VP infants only</i>)

This design permitted both cross-sectional comparisons for group difference at each age-point, and estimations of change over time within-groups on repeated measurements. Tests of association (correlation) were also planned to explore relationships between the various infant measures (attention disengagement task; IBQ-R-VSF; AOSI) at 6/12 months and ADOS-II scores at 30-34 months.

Power Analyses and Sample Size Estimations

Power calculations were carried out using an online power tool (<https://clincalc.com/Stats/SampleSize.aspx>) to estimate the required sample size to detect group differences and associations relating to study hypotheses. Calculations were based on a desired power of 0.8 to detect a difference at an alpha-level of .05.

No study has previously compared a population of VP and full-term infants on the ADOS II during infancy, and as such sample size estimates were based upon a published paper with a lower-risk sample. De Groot, Roeyers and Warreyn (2006) reported findings on the ADOS G (Lord et al., 2000) from a comparison of low-risk preterm infants (31-37 weeks gestation) (N=23) and full-term controls (N=19) at two years of age. From their reported findings it was estimated that a sample size of 16 per group should be sufficiently powered to detect quantitative differences in ASD symptomology at 30-34 month follow-ups. As greater impairment would be expected in a higher risk preterm group (Baron & Rey-Casserly, 2010) this estimate should be considered conservative.

Sample size requirements for the infancy measures at 6 and 12 months (AOSI; attention disengagement task; IBQ-R-VSF) were difficult to estimate given no published data from research with VP samples were available. However, studies using these tasks with *high-risk infant sibling populations* have typically engaged sample sizes of 25-40 participants per group, and

group differences at the required alpha level have been reported (Bedford et al., 2014; Gammer et al., 2015; Zwaigenbaum et al., 2005).

Sample size estimations were completed for detecting associations between AOSI during infancy and ADOS-II scores at follow-up. These estimates were again calculated from findings in a lower risk preterm sample, and as such should be considered conservative. Based on a reported correlation coefficient of .46 between AOSI scores and total scores on the ADOS G (Roncadin et al., 2011), it was estimated that approximately 35 VP infants might be required to detect a significant association between these two measures. Details of a-priori power calculations carried out are shown in Appendix 5.

Prior to commencing the study (i.e. at the time of study proposal), information was available on the total number of infants who had attended assessment visits at 6 and 12 months. It was confirmed that a total of 41 VP and 65 full-term infants had attended 6 month visits, and 38 VP and 43 full-term infants had attended 12 months visits. Therefore, sufficient power to detect group differences, at least on the infancy measures at 6 and 12 months, was anticipated. However, it was not known prior to detailed audit and processing of the existing data how many infants had completed and provided usable data on each measure. This information is reported in the results section.

As this study is part of an *ongoing* longitudinal follow-up study, only a subgroup of infants who have provided data at 6 and 12 months reached the

follow-up age of 30-34 months within the time-scale of the project. As a result it was anticipated that group sample sizes would be lower at 30-34 months than at 6 or 12 months. A total of 23 VP and 14 full term infants reached follow-up age before mid-May 2016 and were invited for assessment.

As this would have resulted in uneven group sizes and possibly insufficient power to detect between-group differences on the ADOS II (see above), a further five full-term control infants were recruited to take part at 30-34 months *only* (i.e. for the ADOS II). There were some disadvantages of this decision. Most notably, it meant that a subset ($N=5$) of infants included in the 30-34 month assessments had not provided any within-subjects longitudinal data, detracting from the uniformity of the overall control sample. Specifically, there were inconsistencies within the control group in terms of recruitment method (outlined below) and age of enrollment in study. However, a larger control sample size at 30-34 months increased the likelihood of achieving a normal distribution in ADOS II scores, helped to balance out group sizes and increased statistical power to detect a significant group difference.

Sample Characteristics

Very Preterm (VP) Group

Inclusion/Exclusion Criteria. The main inclusion criteria for the VP group was birth at ≤ 31 weeks gestation. Exclusion criteria included presence of intraparenchymal haemorrhage on the infant's cranial ultrasound, and/or the presence of any major congenital abnormality (i.e. any abnormality likely

to cause disability using the WHO definition of impairment, disability and handicap or requiring surgery).

Recruitment. All VP infants taking part had previously been admitted to the neonatal intensive care unit (NICU) for specialist neonatal care, after having been born at 22-31 weeks gestation. Recruitment of the preterm sample had taken place on the NICU, whilst infants were admitted for care. A member of the medical research team on the ward was responsible for recruitment of eligible infants.

Sample Overview. In total 41 VP infants took part in the study, though only a proportion of these ($N = 23$) were followed up at 30-34 months. The median gestational age at birth of the *complete* VP sample ($N = 41$) was 26+0 weeks, with a range of 23+3 to 31+3 weeks. Therefore, half of the infants taking part were actually born within the 'extremely preterm' range (i.e. ≤ 26 weeks). This likely reflects the specialist nature of the NICU, which as a tertiary centre with 21 high-dependency cots, provides care for the most medically vulnerable infants.

Correction of Age for Prematurity. For the purpose of follow-up, the age of VP infants was corrected for prematurity (i.e. calculated from due date rather than birth date). Correction of age for prematurity is standard practice in developmental research with preterm infants under 2-3 years to account for expected delays in physical and neurological maturation (Wilson & Craddock, 2004).

Full Term Controls

Inclusion/Exclusion Criteria. All infants in the full term control group had been born between 37 and 42 weeks of gestation, and had birth weight appropriate for their gestational age (i.e. 10-90th percentile). Medical exclusion criteria for the full-term sample were identical to the preterm sample.

Recruitment. Longitudinal full-term infants had been recruited from antenatal classes at the local site. Identification and recruitment of eligible full-term infants was the responsibility of a dedicated study research nurse.

The additional five infants recruited at 30-34 months were recruited from two nurseries local to the service. These nurseries were initially contacted via a formal letter (see Appendix 6). Nurseries that showed interest in taking part were then contacted by telephone to provide more information, and a time to visit the nursery was arranged. A total of three nursery visits were carried out. Information, including a study information sheet (see Appendix 7) was provided to parents of eligible infants aged between 30 and 34 months, which was followed up with a phone call to address questions. A total of seven 30-34 month infants were initially recruited from these three nurseries, but two later cancelled their visits due to other commitments.

Sample Overview. In total 65 full-term infants have taken part in the study. The median gestational age at birth of the *complete* full-term control sample ($N = 65$) was 40+0 weeks, with a range of 37+0 to 42+1 weeks.

Detailed information regarding the clinical and demographic characteristics of the VP and full-term samples was obtained from study questionnaire and medical record information. A summary of this information is provided at the beginning of the Results section (Table 3).

Measures

Details of the early infancy measures used at 6 and 12 months (i.e. AOSI; IBQ-R-VSF; Attention Disengagement Task) are presented first, followed by an overview of the ADOS II and the Bayley III.

The Autism Observation Scale for Infants (AOSI)

Overview. The AOSI (Bryson, 2008) is an 18-item semi-structured direct observational measure designed to measure early behavioural markers associated with ASD in 6–18 month old infants (see Appendix 8). These markers include atypical or delayed social communication behaviours (e.g. social babbling, orientating to name, eye contact, anticipatory social behaviour) and non-social behaviours (e.g. disengagement of visual attention, atypical sensory or motor behaviours), and various aspects of temperament (e.g. reactivity, ease of transitions between activities).

The AOSI assessment comprises a sequence of seven predefined semi-structured activities, which are administered by a trained examiner using a standardized kit of toys: a rattle, a bell, a 'squeaky' toy, blocks, a soft book, a small ball, a rubber duck, a plastic stick and a blanket. Throughout the assessment, the examiner delivers a number of systematic 'presses' or

prompts designed to elicit particular target behaviours (see Table 2). The AOSI assessments at 6 and 12 months were administered by one of three trained examiners, one of whom was the trainee (assessments completed prior to training).

Table 2 Description of the 18 behavioural ‘markers’ coded in the AOSI

Early ASD ‘Marker’	Item Description
Visual tracking	Ability to visually follow a moving object, passing laterally across the midline.
Disengagement of attention	Ability to disengage and move eyes/attention from one of two competing visual stimuli.
Orientation to name	Ability to move head and/or eyes toward and look at the examiner when name is called.
Differential response to facial emotion	Ability to respond differentially through facial, head or other motor movements to a change in the examiner’s facial expression from smiling to a neutral expression (“still face” press)
Anticipatory social response	Ability to anticipate and enjoy social (vs. physical) cause-effect relationships. Ability to smile in response to the examiner’s smile.
Imitation	Ability to reproduce an action produced by the examiner.
Social babbling	Ability to engage in back-and-forth (reciprocal) vocalisations with the examiner.
Eye contact	Ability to consistently establish appropriately sustained eye contact with the examiner.
Reciprocal social smile	Ability to smile in response to the examiner’s smile
Coordination of eye gaze and action	Ability to co-ordinate gaze with actions on objects.
Behavioural reactivity	General responsiveness, including under reactivity and over reactivity, to the activities and toys introduced, and to the examiner’s actions.
Social interest and shared affect	Ease of engagement and interest in activities, and ability to share positive affect with the examiner.
Transitions	Ease and consistency with which toys are relinquished and movement is made from one activity to another.
Motor control	Degree to which motor behaviour is goal-directed, organised and modulated.
Atypical motor behaviour	Presence of developmentally atypical gait, locomotion, motor mannerisms/postures or repetitive motor behaviours.
Atypical sensory behaviour	Presence of developmentally atypical sensory behaviours in any modality (e.g. smelling of toys, staring at hands/shapes/objects, or feeling textures).

Scoring. Using a record sheet, the examiner provides a rating for all 18 marker behaviours. Ratings are made on a scale of 0 to 2 or 3, with higher scores indicating greater atypicality (see Bryson et al. [2008] for details). In

order to aid with reliability of examiners rating, detailed qualitative and quantitative descriptions are provided in the scoring manual. From the examiners ratings, two summary scores are calculated: Total Number of Markers (the number of items on which the infant received a non-zero rating) and Total Score (sum of all 18 item scores).

Psychometric Properties. The AOSI has excellent inter-rater reliability, particularly at 12 and 18 months. At 6 and 12 months respectively, inter-rater reliability for total marker count is .68 and .92 and for total score is .74 and .93. Test–retest reliability is fair to good at 12 months (0.61 for total score), which is respectable in an infant population. For more detail on the AOSI assessment (administration, scoring, reliability and validation) see the published guidelines (Bryson et al., 2008).

Apparatus & Procedure. AOSI administration took place in a dedicated research space within the hospital, when the infant was as alert, happy and responsive as possible. As outlined in the published guidelines, the infant was seated on their parent’s lap at a small table (W70cm x D50cm x H59cm), opposite and facing the examiner. The parent was asked to act as a support/comfort for the infant, but not to assist or guide them in any way for the assessment (“assume the role of an observer”). The assessment is designed to last approximately 15–20 minutes, although administration time is somewhat variable dependent upon the infants’ engagement, temperament, state, and developmental level.

Coding. When the study database audit was carried out, it became apparent that a substantial proportion of assessments had not been coded at the time of the assessment. However, all assessments had been videotaped and this raw data was available to the trainee who has completed research level training on the AOSI. It was therefore decided that the trainee would code all assessments from video to maximize across-participant consistency in coding methods. Examiner ratings, where available, were then used as second ratings to test reliability.

Inter-rater Agreement. Agreement between video and examiner ratings in terms of total score was excellent at both 6 months ($n = 16$, intraclass correlation coefficient = .924) and 12 months ($n = 30$, intraclass correlation coefficient = .962). Agreement on total markers was also excellent at both 6 months ($n = 16$, intraclass correlation coefficient = .852) and at 12 months ($n = 30$, intraclass correlation = .976). Where there was disagreement between coders the examiners ratings were used in analysis.

Outcome Variables. Two dependent variables were obtained for analysis: Total Score and Total Markers (i.e. non-zero ratings).

Attention Disengagement Eye-tracking Task

Overview. The experimental design and stimuli for this task were based upon the task used in Wass, Porayska-Pomsta, & Johnson (2011) and Elsabbagh et al. (2009). The eye-tracking task (including calibration

procedure and stimulus presentation) was written and run using Matlab 2010b.

Stimuli. Stimulus presentation in this task was gaze-contingent (i.e. dependent on the infant looking at targets), meaning that the infant controlled the speed of trial presentation (Wass et al., 2011). For each trial, infants were presented with an animated central fixation stimulus (CFS) at the centre of the screen (a colourful expanding and contracting cartoon ball, subtending 4°). Once the infant was fixated on the CFS, an animated peripheral target (a pulsating cartoon cloud, subtending 3°) was then presented randomly to the left or right of the screen. When the infant's gaze was registered on the peripheral target (PT), one of four short 'reward' animations was shown (e.g. a smiley face or cartoon dog appears, and then springs away with a 'boing' sound).

There were three types of trial in this eye-tracking task: Baseline, Gap, and Overlap. Still-frames of each condition are shown in Figure 2 below. In the 'Baseline' condition, the peripheral target appeared simultaneously with the disappearance of the CFS (i.e. 0ms delay). In the 'Gap' condition, there was a 200ms delay between the disappearance of the CFS and the appearance of the peripheral target. This condition allows *facilitation* effects in visual orienting to be estimated – however, these trials were **not** examined in the present study. On the 'Overlap' trials, the CFS stayed on-screen when the peripheral target appeared, and remained there until the infant shifted their gaze to this peripheral target. The difference between an infant's reaction time

on baseline and overlap trials (i.e. “disengagement effect”) is calculated as a measure of disengagement ability.

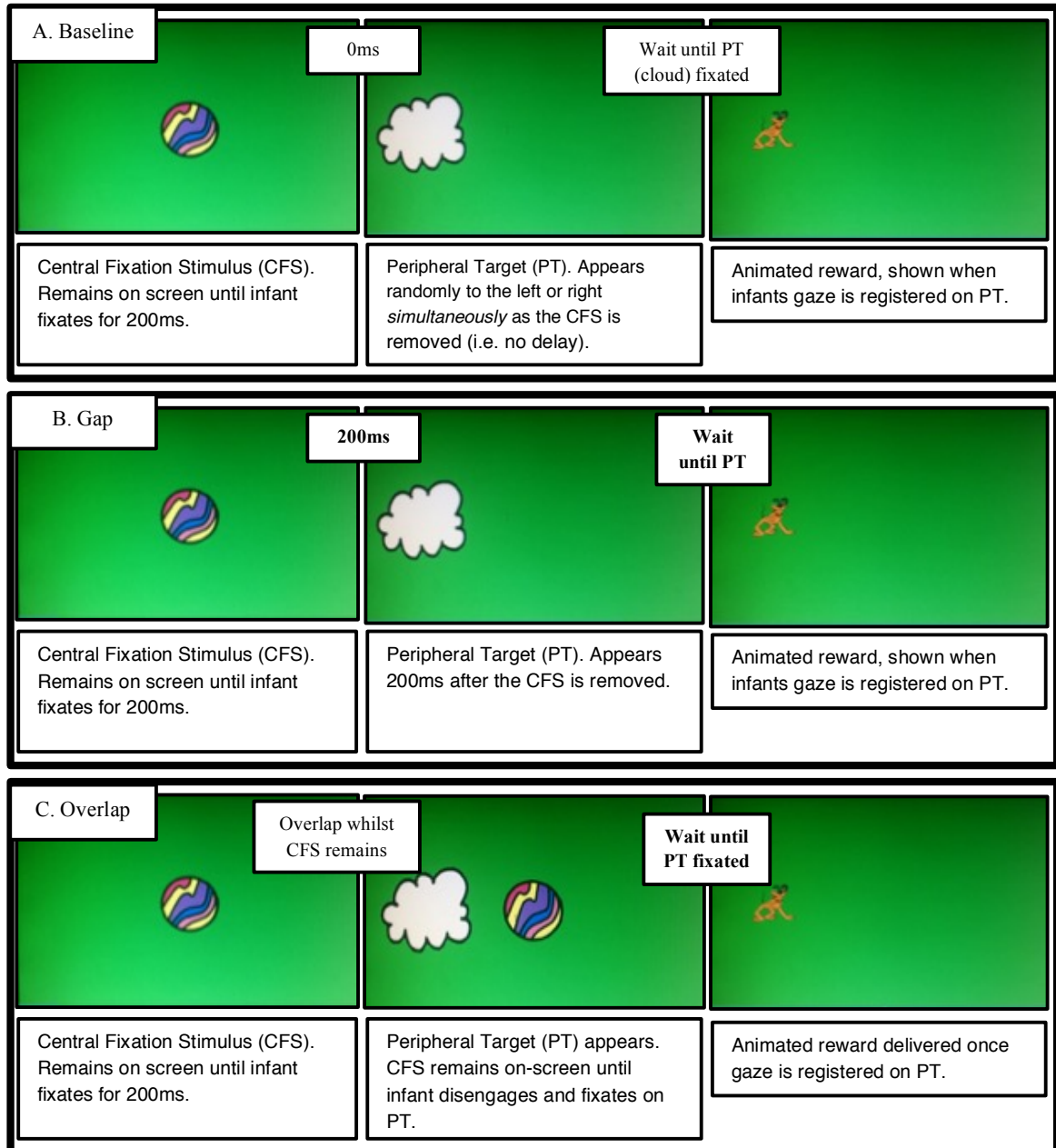


Figure 2 Stimuli (screenshots) of attention disengagement task showing the three trial types: Baseline (A), Gap (B) and Overlap (C)

Stimulus Presentation. The task was presented in blocks of ten trials, within which the three conditions were presented in a pseudo-random order. Between blocks, infants were presented with a brief video animation from Sesame Street™ or ‘In the Night Garden...’ to encourage sustained engagement with the task. If the infant remained engaged and happy, trial presentation continued until 12 usable Baseline and Gap trials, and 16 Overlap trials had been recorded². If the infant became inattentive or fussy before this was achieved, the task was terminated sooner. A minimum of four baseline and four overlap trials was required for inclusion in the sample.

Apparatus & Calibration Procedure. Stimuli were presented on a 40cm x 64cm colour screen. Gaze was measured using a Tobii T60 stand-alone eye-tracker positioned below the screen, and Tobii Studio 3.0.2 software. The equipment was set up in a large plain-paneled booth, to minimise distractions. Infants were seated in a car seat, with their eyes approximately 60cm from the display screen. It was occasionally necessary to seat the child on their mother’s lap if they did not settle in the car seat. Once the infant was settled

² A greater number of overlap trials were required because reaction times were expected to be inflated on these trials (Wass et al., 2011). Increasing the number of trials therefore aimed to minimise anticipated differences in standard error between conditions.

and attending to the screen, the eye-tracker was calibrated using a five-point animated calibration sequence.

Online Filtering of Gaze Data. Communication with the researchers who designed the task at Birkbeck (University of London) confirmed that an 'online' 60ms fixation filter was applied to gaze data to reduce the impact of noise, including minor eye-movements (e.g. micro-saccades or tremors) and random equipment error. Over, Hooge, Vlaskamp, & Erkelens (2007) recommend 60ms as a fixation filter in order to still capture rapid, but genuine, fixations. Gaze data were recorded at 50 Hz, with a spatial resolution of 1" after calibration.

Data Cleaning. In order to obtain the summary variables of interest (i.e. Baseline trial reaction time; Overlap trial reaction time) from the raw eye-tracking data, raw data were processed using a tailored Matlab script. A researcher at Birkbeck (University of London) who had been involved in the development of the task carried out this data cleaning process. The key details of this process were as follows:

- Rectangular areas of interest (AOIs), each subtending 9°, were defined around the CFS and peripheral stimuli.
- Reaction times were calculated as the time elapsed between the appearance of the peripheral target and the infant's gaze *leaving* the CFS.

- A trial was considered valid if a reaction time of between 200ms and 1200ms was recorded (Elsabbagh, Fernandes, et al., 2013; Wass et al., 2011).

Outcome Variables. The three main outcome variables used in analyses were Number of Valid trials, Baseline Reaction Time and ‘Disengagement Effect’ (i.e. Baseline RT subtracted from Overlap RT).

Infant Behaviour Questionnaire Revised Very Short Form (IBQ-R-VSF)

Overview. The IBQ-R-VSF (Gartstein & Rothbart, 2003; Putnam et al., 2014) is a parental report questionnaire used to measure temperament in infants aged between 3 and 12 months. There are numerous conceptions of the construct of Temperament, but the IBQ-R-VSF (and longer forms of the IBQ) is designed to measure individual differences in emotion and motor reactivity, attention, and self-regulation (Rothbart, Ellis, & Posner, 2004), thought to have genetic, psychobiological and environmental underpinnings (Buss & Plomin, 2014). In this study the ‘very short form’ (IBQ-R-VSF) version of the IBQ-R questionnaire was used. This decision was made in order to reduce the burden on parents, given the extensive data collection requirements of the wider research programme, including multiple parent questionnaires.

Scale Description. The IBQ-R-VSF contains 37 items, which contribute towards three broad scales – Surgency (i.e. positive affect), Negative Affect and Effortful Control (i.e. self-regulation) (see Appendix 9). It’s scale structure

and items were derived through factor analysis of the full 14-scale Infant Behaviour Questionnaire-Revised (191 items) (Gartstein & Rothbart, 2003). For each item, the parent is asked to read a description of a behaviour and report the frequency with which the infant exhibited that behaviour over the past seven days. Frequency estimates are circled on a 7-point Likert scale ('Always', 'Almost always', 'More than half the time', 'Half the time', 'Less than half the time', 'Rarely', 'Very rarely', 'Never'). Parents can also select 'Does not apply' on any item.

Psychometric Properties. Reported internal consistency (Chronbach's alpha) on the three IBQ-R-VSF scales are all 0.70-0.73 (Putnam, Ellis, & Rothbart, 2001). Psychometric analysis has demonstrated that the IBQ-R-VSF is strongly correlated with the long-form version of the IBQ-R, and has only slightly lower levels of internal consistency and longitudinal stability (S.P. Putnam et al., 2014).

Data Collection. Parents or caregivers of infants taking part were sent the IBQ-R-VSF when their 6 and 12 month assessment visits were arranged, and either brought the completed questionnaire to the appointment, or completed whilst at the research centre.

Outcome Variables. Three outcome variables were obtained according to the three IBQ-R-VSF scales: Surgency; Negative Affect; Effortful Control. As these represent three separate temperament domains, no overall summary scores are obtained on the IBQ.

Autism Diagnostic Observation Schedule – Second Edition (ADOS-II)

The ADOS-II is a semi-structured, play-based assessment of communication, social interaction, play, and restricted and repetitive behaviours (see Appendix 10). Designed to elicit behaviours directly related to the DSM-V diagnostic criteria for ASD and only administrable by trained researchers or clinicians, the ADOS is considered to represent the 'gold-standard' in ASD assessment (NICE, 2011). Research-determined cut-offs identify the potential diagnosis of ASD, allowing a standardized assessment of ASD symptomatology. The measure has been shown to have good reliability when used by trained clinicians and strong discriminant validity (Lord et al., 2012).

There are five different 'modules' within the ADOS-II, which are selected on the basis of the examinee's language level. Module 1 was deemed to be most appropriate for the majority of infants taking part (i.e. little to no phrase speech). Though some infants at 2 ½ were likely to have developed full phrase speech, and as such may have been able to complete Module 2, it was felt that Module 1 assessment materials were more developmentally appropriate for this age group. Using the same module for all infants taking part also allowed for consistent procedures and scoring structures to be applied, which is preferable in a research context.

Data Collection. ADOS-II assessments were completed by the trainee or one other trained examiner during the 30-34 month visit. All examiners

including the trainee had completed research level training in the administration of the ADOS, including attendance of update courses for the ADOS II. The majority of assessments took place at the research centre in a large distraction free space. It was necessary to complete some assessments during home-visits, as parents were unable to attend the research centre for testing. Administration typically lasted 30-40 minutes.

Coding. Each assessment was coded according to a structured scoring schedule by the examiner immediately after the assessments. As with the AOSI, all assessments were video and audio recorded with parental permission, for code checking. All ADOS assessments were checked from video. Where there was disagreement or uncertainty on specific codes, items were discussed with the examiner and a consensus code was agreed. Consensus codes were used in analyses.

Outcome Variables. Three key summary scores were obtained for each child which represented the main dependent variables on this measure: ADOS-II total score, Social Affect (SA) subscale score and Restricted and Repetitive Behaviours (RRB) subscale score. Total scores were also compared with ADOS II cut-off scores to yield one of two classifications: Autism Spectrum or Non-spectrum.

Identification of Clinical Concern. Five VP infants and one control infant taking part reached diagnostic cut-offs for clinical concern regarding ASD. These cases were notified to the Principal Investigator (an experienced

Neonatologist) who according to clinical judgment then contacted families to offer a referral to local services for further assessment.

Bayley Scales of Infant and Toddler Development - Third Edition (Bayley III)

Scale Description. The Bayley III is a standardised test of motor, language and cognitive development for infants aged from 1 to 42 months of age (Bayley, 2006). It is widely used in Europe and the US, both in clinical practice and research. The Bayley III contains three subscales which yield three summary scores: Cognitive; Language (Expressive and Receptive) and Motor (Fine & Gross motor). Unlike its predecessor, the Bayley III does *not* provide an overall summary score for all three subscales.

Administration & Scoring. The assessment is carried out by a trained examiner and comprises a series of predefined tasks which are presented to the infant. Each item is scored according to a pass or fail rule. The infants age at assessment, plus discontinuation and reverse rules, determine the items completed by the infant. Raw scores for each scale can be used to derive composite and scaled scores which enable comparison to the normative sample. Notably, the Bayley III has come under some scrutiny for providing Cognitive normative scores approximately seven points higher than its predecessor (Bayley II) (Robertson, Hendson, Biggs, & Acton, 2010). This should be taken into account when interpreting infants' absolute scores on the Bayley III in this study.

Collation of Bayley III scores. All infants in the VP group had completed the Bayley III assessment as part of their routine outpatient developmental follow-up at 12 months and 24 months. These scores were obtained with permission from parents from infants' hospital records (in line with Ethics approval). The Bayley III assessment had also been completed with full-term control infants at their 12 month follow-up visit. These scores were obtained from the wider project database.

Outcome Variables. Cognitive, Motor and Language composite summary scores were calculated.

General Procedure

As explained previously, this project constituted two broad procedural components. The first was the audit and processing of early infancy data already collected as part of a wider study of development after preterm birth. The second component comprised the longitudinal follow-up of these infants at 30-34 months to measure ASD outcomes.

Collection and Audit of Existing Infancy Data. Existing data from the 6 and 12 month assessments was collated from secure electronic and paper records. A substantial amount of data processing was required to prepare existing data for analysis. All available questionnaire data (i.e. demographics and IBQ-R-VSF) had been inputted into the research programme database, but was double checked from paper records to ensure accuracy. Eye-tracking data for the attention disengagement task was still in raw form (i.e.

unprocessed), and summary variable data was extracted in collaboration with researchers who had developed the task at Birkbeck (University of London). All AOSI videos were coded from video recordings and individual item and summary scores (Total scores and Total Markers) were inputted into the study database.

Collection of New Data (ADOS II) at 30-34 months. The majority of ADOS II assessments took place in a dedicated research space within the hospital. However, it was necessary to complete home visits for four VP infants and three full-term controls. Parents (or legal guardians) of infants enrolled in the longitudinal study had already provided their written consent to take part, including follow-up at 2.5 years (see Appendix 11). However, at the start of appointments parents were given the opportunity to look over information sheets again and ask any questions about taking part. Parents of infants taking part at 30-34 months only provided written consent prior to commencing data collection (see Appendix 12).

Collection of data used in this study took place alongside data collection for various other tasks *not* included in this report. Families' visits to the Babylab typically lasted around 3-4 hours, but a large degree of flexibility was required to try to maximise infants' cooperation. For instance, it was often necessary to provide breaks for feeding, changes or naps. On occasion, it was not possible to complete all study tasks, due to poor infant compliance. Where possible, assessments were rearranged to pursue missing data, either at the research centre or as a home-visit.

Details of missing data and overall retention rates are shown in the Results section.

Data Analysis Procedure

Data Exploration

Prior to implementation of any inferential statistics, data distributions of all continuous dependent variables were first explored to identify any outliers and to determine whether normality could be assumed. This included visual examination of box plots and histograms, and carrying out numerical significance tests of skewness and kurtosis in SPSS. Where possible, statistical transformations were performed to improve the normality of distributions. The only dependent variable that remained non-normal following transformation was ADOS subscale scores (i.e. social affect/repetitive behavior). A series of bivariate comparisons were also carried out to check VP/full-term group matching on various demographic indices (gender; SES; bilingual home).

Cross-Sectional Comparisons

As ADOS II subscale scores were not normally distributed, non-parametric tests were used for the examination of between-group differences in Social Affect and Repetitive behaviours at 30-34 months. Non-parametric tests were also used for estimation of group differences in categorical screening outcomes on the AOSI and ADOS II, using established scale cut-offs as binary dependent variables. Between-group comparisons on all other continuous dependent variables were approached using parametric tests (t-

tests/analysis of variance). In recognition of the number of planned comparisons being carried out, Bonferroni corrections were applied throughout – the details of which are outlined in the results section.

As the two groups were shown to differ in terms of gender distributions and estimated social economic status, further ‘checks’ were carried out using analysis of covariance and correlational analyses to estimate the possible influence of these variables upon any significant between-groups differences observed.

An additional exploratory hierarchical regression analysis was implemented to estimate the influence of developmental delays on AOSI group differences.

Longitudinal Associations

To address longitudinal associations between early infancy measures and ASD outcomes at 30-34 months in the preterm group, it had been anticipated that a hypothesis driven hierarchical regression approach would be used. However, multiple regression was not justified given the within-subjects VP sample size obtained. As a result, correlational analyses and partial correlations were used as a preliminary exploration of longitudinal associations between early infancy measures, ADOS II scores and Bayley III summary scores.

Service User Involvement

Feedback from parents regarding their experience of taking part was collected routinely at assessment visits using a brief questionnaire (see Appendix 13).

This questionnaire had been used at previous (6m/12m) follow up visits, and parent feedback was used to inform planning of procedures at the 30-34 month follow-up. Feedback was broadly positive, but parents had some helpful suggestions regarding timings of assessments. For instance, one parent felt that offering weekend appointments would be particularly beneficial for 30-34 month visits to accommodate parents work commitments and child-care needs. At the end of study visits, parents were also asked if and how they would like to be informed of the study outcomes.

Chapter 3: Results

The following chapter is broadly divided into three sections. The first section will provide an overview of sample characteristics and adequacy of matching between groups. The second section will address cross sectional comparisons between the two groups, at 6 months, 12 months and at 30-34 month follow up. The third section will consider longitudinal associations between the early infancy measures, and ASD symptomatology (ADOS II) at 2.5 years.

Sample Characteristics

Table 3 below summarises the total number of VP and full-term infants who attended visits at each time point, as well as the number of those infants who actually contributed *usable* data for each individual measure. This highlights the variation in sample sizes between measures and follow-ups in this longitudinal data-set.

Table 3 *Total Number of Infants participating in the study at 6, 12 and 30-34 months with N's for Individual Measures*

		VP Group	Full Term Group
6 months	Attended Visit (i.e. Total N)	41	65
	IBQ-R-VSF	36	54
	Attention Disengagement Task	24	34
	AOSI	33	39
12 months	Attended Visit (i.e. Total N)	38	43
	IBQ-R-VSF	35	39
	Attention Disengagement Task	23	31
	AOSI	29	31
30-34 months	Attended Visit (i.e. Total N)	23	19
	ADOS	22	18* (*13 Longitudinal)

The most notable reason for sample size discrepancies across time-points is that data are from an ongoing developmental follow-up study, and some infants have not yet reached 12 or 30 months of age. As such, the largest sample sizes were obtained for 6m cross-sectional comparisons. There was also, however, a substantial amount of missing data from specific assessments, particularly within the existing infancy (i.e. 6 and 12 month) data-sets. Reference to archived data-collection records indicated that the most frequent reason for non-valid or missing data was infant non-compliance (e.g. due to fussiness/fatigue).

There was also some missing data due to missed follow-up visits or participant drop-out, with retention rates from 6 to 30-34 months of 82% in the VP group and 74% in the term group. Retention rates in the full-term control group in particular were affected by delays in obtaining ethics clearance for follow-up at 30-34 months, which led to six infants missing follow-up at this age-bracket (i.e. they were too old by the point of ethics approval). Across the groups, the most frequent reasons for study drop out were moving out of the area, change in family circumstances or inability to contact. Two additional infants in the preterm group dropped out due to ill-health.

In order to characterise the whole sample, demographic summary information for the total VP and full-term samples (i.e. all infants who provided data at any time point) is shown in Table 4. Information regarding infants length of stay in specialist care was obtained from medical history information, and was collated by the study research nurse.

Table 4 Overall sample characteristic of the VP and full-term control groups, including all infants contributing data at 6m, 12m and/or 30-34 months

	Very Preterm Sample		Full-Term Control Sample	
<i>N</i>	41		65	
Gender	<i>Male (N)</i>	27	<i>Male (N)</i>	30
	<i>Female (N)</i>	14	<i>Female (N)</i>	35
Gestational Age at Birth (Days)	Mean (SD)	184.12 (13.19)	Mean (SD)	281.2 (9.35)
	Median	182 (i.e. 26+0)	Median	282 (i.e. 40+1)
	Range	165-221 days (i.e. 23+3-31+3)	Range	259-295 days (i.e. 37+0-42+1)
Birthweight (g)	Mean (SD)	821.0 (220.88)	Mean (SD)	3476.83 (481.51)
	Median	774.0	Median	3400
	Range	460g-1521g	Range	2500g-4570g
No Days in Tertiary Care Post-Birth	Mean (SD)	86.2 (47.81)		
	Median	76		
	Range	14-249 days		
SES Indicator - Parent Education Level (Scale 1-8)	Mean (SD)	6.68 (1.88)	Mean (SD)	7.31 (1.08)
Multiple Birth (percentage)	34%		10%	
Bilingual Exposure	43%		36%	

Demographic Matching

A series of between group comparisons was carried out to ascertain the adequacy of matching on demographic matching variables. To ensure a conservative approach to matching, no Bonferroni corrections were applied. Chi square tests indicated that there was a significant difference between the groups in terms of gender distribution ($\chi^2(1) = 3.925, p = .048$). An independent samples t-test also indicated that there was a marginal difference between the two groups in terms of mean parental education level – an estimator of social economic status (SES) ($t(57.068) = -1.931, p = .058$). Bilingual exposure could be considered equivalent across the groups ($\chi^2(1) = .512, p = .474$).

It was also noted that the proportion of multiple births was greater in the very preterm group compared to the full-term sample ($\chi^2(1) = 10.196, p = .001$), which is characteristic of VP populations (Goldenberg et al., 2008).

Cross Sectional Comparisons

A series of between-subjects analyses were carried out to compare the two groups on infancy variables (AOSI; Disengagement of Attention; IBQ-R-VSF) at 6 and 12 months, and on ADOS II outcomes at 30-34 months. Prior to implementation of these analyses, the data were explored to determine whether assumptions of normality had been met. This included completing statistical significance tests of skewness and kurtosis in SPSS, and visual reference to histograms of data distribution (Chen, Ender, Mitchell, & Wells, 2003). Where problematic skew or kurtosis was observed, various

transformations (log; square; square root; reflection and adding a constant) were attempted. Outcomes of transformations are reported. Boxplots were also examined to highlight any univariate outliers. Field (2013) advises that observations greater than three standard deviations away from the mean should be considered problematic.

Exploration of demographic variables indicated that there were notable differences between the two groups in terms of gender distributions and SES. Therefore, where significant group differences were observed, bivariate comparisons were followed up with further ‘post-hoc’ tests to examine the possible influence of these demographic indices.

Autism Observation Scale for Infants

Means and standard deviations of AOSI Total Scores and Total Markers are shown in Table 5. Examination of histograms and normality statistics indicated that the distribution of AOSI total scores and markers could be considered normal in both groups. No outliers were identified.

Table 5 Descriptive Statistics for AOSI Total Scores and Total Markers at 6 and 12 months for VP and full-term infants

		6 Months		12 months	
		VP (N = 33)	full-term (N=39)	VP (N = 29)	full-term (N=31)
Total Score	Mean	10.79	7.62	7.72	5.00
	SD	4.36	3.75	3.61	3.22
Total Markers	Mean	6.94	5.31	5.59	2.03
	SD	2.47	2.20	3.84	2.22

Bivariate Analysis of AOSI Continuous Scores.

T-tests. A series of independent samples t-tests were carried out comparing Total scores and Total markers across the two groups at 6 and 12 months. A Bonferonni correction was applied to adjust for multiple comparisons. As Total markers and Total scores are closely inter-related, adjustment was made only for the multiple time-points (i.e. .05/2).

At 6m, VP infants obtained higher AOSI Total scores than full-term controls ($t(70) = 3.320, p=0.001$) as well as significantly more Total markers ($t(70) = 2.960, p=0.003$). At the 12m follow-up, the VP group again received significantly higher Total scores ($t(58) = 3.094, p=0.003$) and Total markers ($t(58) = 3.175, p=0.002$) than the full-term control group.

Mixed ANOVA. To address within-subjects change amongst the subset of infants who had completed the AOSI at both 6 and 12 months, a mixed ANOVA was carried out, with group (VP; full-term) as the between-group factor and time-point (6m; 12m) as the within groups factor. To limit the number of planned comparisons, only AOSI *Total* scores were included in this analysis as Total scores and Markers showed a similar pattern across the two groups in bivariate analyses.

As expected from bivariate analyses at 6/12 months, there was a main effect of group; VP infants in this subgroup showed higher Total scores ($M= 9.365, SE = .536$) than full term controls ($M = 6.769, SE = .536$) across the two follow-ups ($f(1, 50) = 11.749, p=.001, \eta p2 = .190$). There was also a

significant main effect of time-point, with infants generally showing greater 'impairment' on the scale at 6 months ($M = 9.385$, $SE = .557$) than at 12 months ($M = 6.750$, $SE = .474$), ($f(1, 50) = 13.991$, $p < .001$, $\eta^2 = .219$). There was no significant interaction effect, indicating that the difference between AOSI Total scores in the two groups was equivalent at 6 and 12 months ($p = .935$) (see Figure 3).

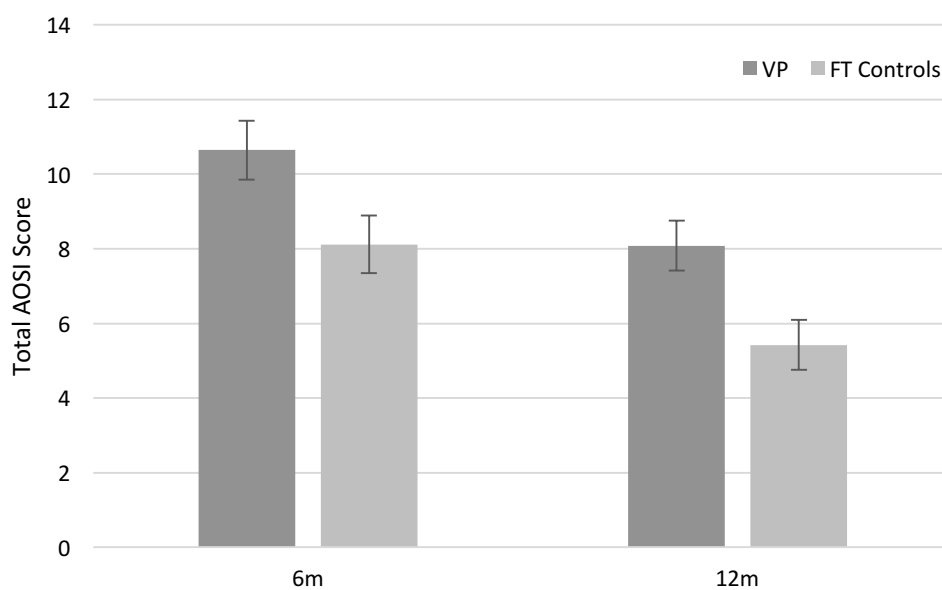


Figure 3 Graph showing significant main effects of group and time-point for AOSI scores

Estimating the Influence of Gender and SES

As gender and SES³ were not sufficiently matched across the two groups, it was important to consider the role these variables may have played in observed group differences on the AOSI.

Selection of Appropriate Tests. Miller and Chapman (2001) argue that whilst analysis of covariance (ANCOVA) should not be used to control for covariates that show substantive differences between groups, it *can* be appropriate where there is good reason to believe these differences arose by chance. In the case of the uneven gender distribution between the two groups (i.e. more boys the preterm group), though there is evidence of a slight over-representation of males in preterm populations, reported odds ratios in whole-population samples are estimated to be only marginally above 1 (1.06) (Zeitlin, Ancel, Larroque, Kaminski, & the Epipage Group, 2004). Such a large over-representation of boys in the present study (approximately 2:1) is therefore unprecedented, and is likely to have arisen by chance in this relatively small sample. There was no reason to believe that the high proportion of males in the sample had any meaningful relationship with the site, sampling or recruitment process. Miller and Chapman also state that it can be helpful to conduct ANCOVAs in order to understand patterns of shared

³ SES estimated from reported parental education level, rated on a scale of 1-9 (see Appendix 14). Where ratings were provided for both parents, the mean was calculated. Where only one parents education level was available, this rating was used.

variance in a data set, whilst taking care in how these 'effects' are reported. Therefore ANCOVA was deemed to be an appropriate tool to explore any *possible* influence of gender on the results.

ANCOVA was deemed *not* to be suitable for estimating possible influences of SES within the sample. Firstly, associations between lower social economic status and increased rates of spontaneous and indicated preterm birth are known to exist (Joseph et al., 2007). Secondly, inherent differences in sampling and recruitment procedures across the two groups may have affected between-group differences in SES. Where the covariate and grouping variable cannot be considered independent, the findings of ANCOVA become very difficult to interpret. As a result, Pearsons correlations were carried out to highlight any associations between SES and AOSI scores that *may* be influential in group differences.

Gender (ANCOVA). Two one-way ANCOVAs were conducted to determine whether AOSI total scores varied between the VP and full-term groups *after* controlling for infants' gender at 6 and 12 months. Levene's test indicated adequate homogeneity of variance at both 6 months ($p = .789$) and 12 months ($p = .668$). At 6 months, there was a main effect of birth-status ($F(1, 68) = 10.023, p = .002, \eta_p^2 = .128$), with infants in the VP group scoring higher than those in the full-term group. There was no main effect of gender ($F(1, 68) = .027, p = .870, \eta_p^2 = .000$), and no significant interaction between birth-status and gender ($F(1, 68) = .065, p = .799, \eta_p^2 = .001$).

A similar pattern was observed at 12 months; there was a main effect of birth-status ($F(1, 56) = 8.764, p = .004, \eta_p^2 = .135$), with infants in the VP group scoring higher than those in the full-term group. There was no significant effect of gender ($F(1, 68) = .113, p = .738, \eta_p^2 = .002$), or interaction between birth-status and gender ($F(1, 68) = .006, p = .937, \eta_p^2 = .000$). Together this suggests that elevated AOSI scores in the preterm group relative to controls are *unlikely* to be explained by gender differences.

Social Economic Status. Pearson's product moment correlations were carried out to explore possible associations between estimated SES (parental education level) and AOSI scores. These showed no significant correlation between SES and AOSI scores at 6 months ($r(71) = -.026, p = .829$) or at 12 months ($r(59) = -.234, p = .075$), indicating that SES is unlikely to be a notable influence upon observed group differences on the AOSI.

Hierarchical Regression

In line with research hypotheses, a further analysis was carried out to estimate the impact of developmental delays in the VP group in observed between-group differences on the AOSI. As the Bayley III had only been completed with infants at 12 months, only 12 month AOSI scores were considered.

Composite scores obtained from existing Bayley III databases are summarised in Table 6. It is notable that language and motor composite scores were not available for all full-term infants, and therefore the number of infants contributing scores towards each scale are presented.

Table 6 Composite Scores for Cognitive, Language and Motor Subscales on the Bayley III

	Very Preterm Infants M (SD)	Full Term Control Infants: M (SD)
Cognitive***	97.76 (9.60)	107.44 (12.02)
	(N=29)	(N=41)
Language	99.10 (18.25)	100.24 (11.17)
	(N=29)	(N=33)
Motor**	92.38 (13.63)	101.46 (12.94)
	(N=29)	(N=35)

*** $p < .001$

** $p < .01$

A series of independent t-tests were carried out to examine group differences in motor, cognitive and language functioning. Exploration of distributions indicated that assumptions of normality were met, and no problematic outliers were identified. Composite scores in the VP group were lower than the full-term group on the cognitive ($t(67.167) = -3.710, p < .001$) and motor scale ($t(62) = -2.727, p = .009$), but language scores were equivalent across the two groups ($t(45.188) = -.292, p = .772$).⁴

⁴ As composite scores are standardized scores (Mean = 100; SD = 15; Range 40-160) based on a normative sample (Bayley, 2006), these differences appear to reflect motor *delay* in the VP group, but above average cognitive ability in the full-term group. Though this may reflect an above control average sample, it is likely that this observation is an artefact of the Bayley III cognitive scale which has been demonstrated to *overestimate* normative cognitive ability by approximately 7 points compared to previous versions of the Bayley and other scales of infant development (Anderson, De Luca, Hutchinson, Roberts, & Doyle, 2010; Picciolini et al., 2015).

A hierarchical regression approach was used to investigate the extent to which discrepancies in language and motor functioning levels accounted for variance in AOSI Total scores. As hierarchical regression allows the researcher to enter independent variables into the model in a theoretically informed order, it can be used to estimate the contribution of an explanatory variable, after the effects of a control variable have been 'partialled out' (Field, 2013).

First, a simple regression was performed with 12m AOSI Total score as the dependent variable, and birth status as a single predictor (**Model A**). Two hierarchical regressions were subsequently carried out:

Model B: 12 month Cognitive Composite scores (Step I) and birth status (categorical variable: VP; full-term) (Step II)

Model C: 12 month Motor Composite scores (Step I) and birth status (categorical variable: VP; full-term) (Step II).

Checking Assumptions. Histograms and P-P plots were consulted to check normality of residuals for all models. No indications of notable heteroscedasticity (i.e. funneling of data points) or non-linearity (curvature) were detected using plots of standardized residuals against standardized predicted values. Partial plots were examined for obvious outliers (that might have undue influence on a regression coefficients), and no problematic residuals were observed. VIF values for multiple regression models were all well below 10, which indicates no difficulties with multicollinearity (Field,

2013). Power to detect significant associations was estimated for Model A and Model B using a ‘retrospective’ power calculator online (Soper, 2016). Using observed R^2 values, sample size, a required alpha of .05 and number of predictors ($N=2$), power estimates of .78 and .75 were obtained, where it is conventionally stated that a power of .8 is desirable (Ellis, 2010).

Models.

Model A (see Table 7) confirmed that birth status (VP; Full term) was a significant predictor of 12m AOSI Total Score ($F(1, 59)=9.571, p=.003; R^2=.142$, adjusted $R^2=.127$. Reference to the variable coding (VP = 1; Full-term = 2) and the negative beta value indicates that prematurity was associated with greater impairment on the AOSI, as expected.

Table 7 *Model A: Simple linear regression including birth status only*

	<i>b</i>	SEB	β	<i>P</i>
Step 1				
Constant	10.45	1.41		$p=.000$
Birth Status	-2.72	.88	-.38	$p=.003$

Model B (see Table 8) showed that Bayley cognitive score did not explain a significant amount of variance in AOSI score ($F(1, 52)=.134, p=.716; R^2=.003$, adjusted $R^2=-.017$). After controlling for cognitive development, birth status also remained a significant predictor of 12m AOSI score ($F(1, 51) = 8.888, p=.004; R^2=.151$, adjusted $R^2=.117$).

Table 8 Model B: Hierarchical linear model including Bayley cognitive scores (Step 1) and birth status (Step II)

	<i>b</i>	SEB	β	<i>p</i>
Step 1				
Constant	7.80	4.46		<i>p</i> =.086
12m Bayley Cognitive	-.016	.043	-.05	<i>p</i> =.716
Step 2				
Constant	7.40	4.15		<i>p</i> =.081
12m Bayley Cognitive	.36	.04	.17	<i>p</i> =.413
Birth Status	-3.13	1.05	-.42	<i>p</i> =.004

Similarly, Model C showed that Bayley motor score did not predict AOSI score ($F(1, 49) = .102, p = .750; R^2 = .002$, adjusted $R^2 = -.018$) (see Table 9).

Furthermore, after controlling for motor development, birth status was still a significant predictor of 12m AOSI score ($F(1, 48) = 8.308, p = .006; R^2 = .149$, adjusted $R^2 = .114$).

Table 9 Model C: Hierarchical linear model including Bayley Motor scores (Step 1) and birth status (Step II)

	<i>b</i>	SEB	β	<i>p</i>
Step 1				
Constant	7.39	3.78		<i>p</i> =.056
12m Bayley Motor	-.01	.04	-.05	<i>p</i> =.750
Step 2				
Constant	9.15	3.57		<i>p</i> =.014
12m Bayley Motor	.02	.04	.07	<i>p</i> =.633
Birth Status	-3.03	1.05	-.40	<i>p</i> =.006

Hierarchical Regression Summary. These models indicated that though differences in motor and cognitive functioning existed between the two groups, neither independently predicted AOSI score at 12 months – and after their influence was partialled out, prematurity continued to be a significant predictor of AOSI scores.

Comparison of AOSI Positive Screening Rates

The developers of the AOSI propose a cut-off criterion of ≥ 7 markers for identification of infants at increased risk for later diagnosis with ASD (Zwaigenbaum et al., 2005). This cut-off has been used in previous research to identify the highest risk infants or ‘positive screens’ (e.g. Yaari et al., 2016). Absolute numbers and proportions (percentage) of infants falling above this cut-off criterion at 6 and 12 months are shown in Table 10.

Table 10 *AOSI positive screening rates (7+ markers) at 6 and 12 months*

	Very Preterm	Full-Term
6 months	<i>N</i> = 33	<i>N</i> = 38
Positive Screen Count	18	11
Percentage	54.5%	28.9%
12 months	<i>N</i> = 29	<i>N</i> = 31
Positive Screen Count	9	3
Percentage	31%	9.7%

Two chi square analyses (Pearson Chi square) were carried out to compare rates of positive screening in the VP and full-term groups at 6 and 12 months. VP infants were significantly more likely to screen positive than full-term controls at 6 months ($\chi^2(1) = 64.79, p = .029$) and at 12 months ($\chi^2(1) = 4.271, p = .039$).

Attention Disengagement

Data Exploration. At the 6m visit, a total of 24 VP infants and 34 full-term infants completed the disengagement of attention task. An additional 11 infants completed the task (8 VP; 3 full-term), but had to be excluded as they completed insufficient trials (a minimum of four valid trials per condition was required for inclusion). At 12 months, a total of 23 VP infants and 31 full-term controls provided valid data. A further 4 VP infants and 9 full-term who completed the task were excluded from analysis as trial number criterion were not met.

Table 11 shows the means and standard deviations for Baseline Reaction Time (RT), “Disengagement Effect” (i.e. Mean Overlap trial RT minus Mean Baseline Trial RT) and total number of valid trials at 6 months and 12 months. Though the main independent variable of interest is Disengagement effect, Baseline RT was included as it provides an estimate of infants’ visual orienting ability. Total Number of trials was of interest as it provides a crude measure of infants’ general level of engagement in the task (i.e. sustained attention).

At 6 months, Baseline trial RTs in the full-term group showed positive skew according to histograms and statistical significance tests ($z = 3.355$, $p < .001$). The distribution of Total Valid Trials in the full-term group was also positively skewed at 12 months ($z = 4.164$, $p < .001$). After application of a log transformation in SPSS, distributions of both variables could be considered normal according to criterion for skewness z-scores (i.e. < 3.29 or $p > .001$; MacLeod, 2015). Therefore log transformed data were used in between groups analyses for Valid trials and Baseline trial RT. All other distributions met assumptions of normality without transformation. No problematic outliers were identified.

Table 11 *Attention Disengagement Summary Scores at 6 and 12 months*

		6 months		12 months	
		Very Preterm ($n = 24$)	Full-Term ($n = 34$)	Very-Preterm ($n = 23$)	Full-Term ($n = 31$)
Baseline RT (ms)	Mean	383.81	293.70	384.18	378.70
	SD	60.527	36.443	33.312	45.508
Disengagement Effect (ms)	Mean	284.03	293.15	316.77	323.87
	SD	128.749	130.384	118.702	112.105
Number valid Trials	Mean	42.21	43.00	38.83	44.84
	SD	9.255	5.810	9.079	11.267

Simple Bivariate Comparisons. In line with research hypotheses, a series of three independent samples t-tests were carried out to explore

between group differences on Baseline RT, disengagement effects and Total Number of Valid trials, at both 6 and 12 months. A Bonferonni correction ($\alpha/3$) was applied to adjust for multiple comparisons at each age, which adjusted the alpha-level for significant group difference to $p < .016$.

The bivariate comparisons indicated that infants in the two groups completed an equivalent number of valid trials 6 months ($t(56) = -.400, p = .691$) and at 12 months ($t(52) = -2.270, p = .027$). However, it is notable that the latter effect was 'marginal' (with Bonferonni correction). There was also no significant difference between VP infants and full-term control infants in terms of disengagement effect at both 6 months ($t(56) = -.264, p = .793$) and 12 months ($t(52) = -.224, p = .823$). The two groups did differ in terms of Baseline RT at 6 months ($t(56) = 7.512, p < .001$), but not at 12 months ($t(52) = .488, p = .627$), with VP infants showing slower visual orienting than full-term infants.

Within-subjects Effects. As with the AOSI, a subset of infants had completed this task at *both* 6 and 12 months (VP $N = 18$; full-term $N = 19$). A mixed ANOVA was carried out within this subgroup, with group (VP; Full-term) as the between-group factor and time-point (6m; 12m) as the within groups factor. Only *Disengagement Effect* was included in this repeated measures analysis as it is key to cross-sectional research hypotheses.

There was no main effect of group ($F(1, 35) = 1.190, p = .283, \eta^2 = .033$) or time-point ($F(1, 35) = 1.828, p = .185, \eta^2 = .050$), and no significant group x time-point interaction ($F(1, 50) = 1.096, p = .302, \eta^2 = .030$). Though it is

important to note that limited power may have contributed towards failure to detect these small-moderate effects (Pallant, 2007).

Temperament (IBQ-R-VSF)

Parent-reported infant temperament (IBQ-R-VSF) questionnaires were collated from a total of 36 VP and 54 full-term control infants at 6 months, and 35 VP and 39 full-term infants at 12 months. Descriptive statistics for each IBQ-R-VSF domain (surgency; negative affect; effortful control) are summarised in Table 12.

Table 12 *IBQ-R-VSF Total and Subscale Scores at 6 and 12 months*

		6 months		12 months	
		Very Preterm (n= 36)	Full-term Control (n= 54)	Very Preterm (n = 35)	Full-term Control (n= 39)
Surgency Subscale	Mean	4.67	3.39	5.46	5.26
	SD	.89	.83	.62	.76
Negative Affect	Mean	3.39	3.42	4.28	3.78
	SD	.83	.75	1.02	.79
Effortful Control	Mean	4.65	3.41	5.33	5.16
	SD	.82	.78	.60	.79

Exploration of data distributions indicated significant negative skew in 12 month *Surgency* scores in the preterm group ($z = -4.547, p < .001$). After reflecting scores (i.e. multiplying by -1) and adding a constant to translate

negative skew into positive skew (Tabachnick & Fidell, 2007⁵), a logarithmic transformation was applied in SPSS. Visual examination of the histogram and statistical indicators of normality suggested that a normal distribution was achieved (z-score (skewness) = 1.404, $p > .01$; z-score (kurtosis) = .689, $p > .01$). All other dependent variables from the IBQ-R-VSF could be considered normally distributed. One possible outlier was observed in the VP group on the *Negative Affect* scale. In order to retain this data point (and thus the representativeness of the clinical sample) it was 'trimmed' (Carey, Carey, Maisto, & Henson, 2004) to be within 3 SD of the mean (Tabachnick, Fidell, & Osterlind, 2007).

A series of independent samples t-tests were carried out to compare IBQ-R-VSF domain summary scores across the two groups. A Bonferroni correction was again applied to adjust for multiple comparisons at each age-point (.05/3), resulting in an adjusted alpha-level of .017.

A marginal between groups difference was observed on the *Negative Affect* scale ($t(72) = 2.382$, $p = .022$) (see Figure 4). All other between-groups comparisons were non-significant ($p > .05$).

⁵ Transformation of negative skew is more challenging as direct application of logarithmic or square root transformations further exaggerates skewness. Tabachnik and Fidell (2007) advise to translate negative skew into positive skew, thus enabling the use of log, square root or fractional transformations.

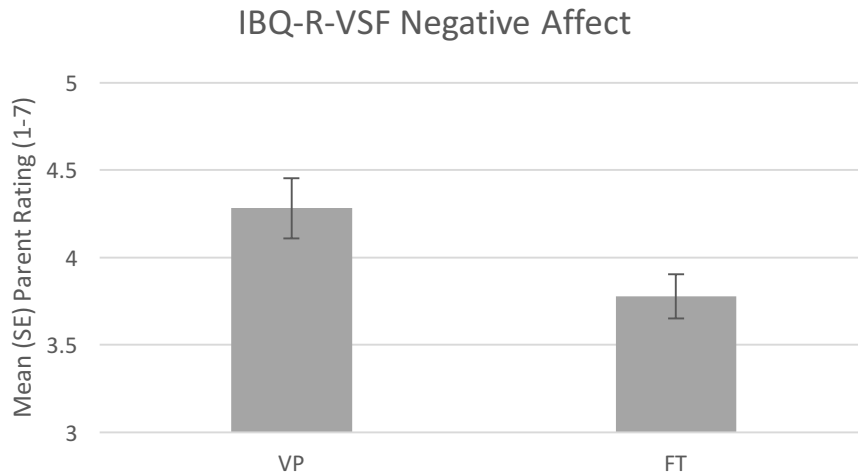


Figure 4 Illustration of marginally significant difference ($p=.022$) in parent reported Negative Affect at 12 months of age

Estimating the Influence of Possible Confounds: Gender and SES

A one-way ANCOVA was carried out to determine whether negative affect ratings varied between the VP and full-term groups after controlling for infants' gender. Levene's test indicated that homogeneity of group variance could be assumed ($p = .086$). The main effect of birth-status was significant ($F(1, 70) = 7.617, p = .007, \eta p2 = .098$), with parents of VP infants reporting higher levels of negative affect. There was also a marginally significant effect of gender ($F(1, 70) = 3.760, p = .057, \eta p2 = .051$), with more negative affect reported in the female infants taking part. The interaction effect (gender x birth status) was non-significant ($F(1, 68) = .041, p = .841, \eta p2 = .001$), suggesting that the effects of gender were equivalent across the two groups.

Overall, this indicates that group differences on the negative affect scale of the IBQ-R-VSF remained significant after controlling for variance associated

with gender; though the male gender bias in the VP sample may have actually moderated (i.e. reduced) observed group differences.

Using Pearson's product moment correlations, negative affect scores on the IBQ-R-VSF were found to be unrelated to SES ($r(73) = -.049, p = .684$). It is thus perhaps unlikely that SES played an influential role in observed (marginal) group differences in *Negative Affect*.

Autism Diagnostic Observation Schedule (ADOS II)

A total of 23 VP infants and 19 full-term infants were invited to complete the ADOS II at 30-34 months, including five full-term controls who had *not* taken part in the full longitudinal study. The assessment was successfully completed with a total of 22 VP infants and 18 full-term control infants⁶. Sample characteristics of infants included in this analysis are summarised in Table 13.

⁶ One VP assessment and one full-term assessment were incomplete (terminated early) owing to infant non-compliance.

Table 13 *Sample characteristics for VP and full-term infants who completed the ADOS II at 30-34 months*

	VP (N = 22)		Full-Term Control (N=18)	
Age at Visit (Mean/SD)	940 days (27.61)		958 days (58.79)	
Gender	<i>Male (N)</i>	15	<i>Male (N)</i>	8
	<i>Female (N)</i>	7	<i>Female (N)</i>	10
Gestational Age at Birth (Days)	Mean (SD)	184.59 (13.34)	Mean (SD)	280.5 (11.163)
Birthweight (g)	Mean (SD)	857.0 (241.95)	Mean (SD)	3465.56 (492.09)
SES Indicator - Parent Education Level (Scale 1-8)	Mean (SD)	6.83 (1.66)	Mean (SD)	7.22 (1.12)
Bilingual Exposure	41%		39%	

To check between-group matching in this smaller sample, a series of bivariate comparisons were carried out comparing gender distribution, age at visit, parent education level and bilingual exposure. Chi square tests indicated that there was no significant difference between the groups in terms of gender distribution ($\chi^2(1) = 2.283, p = .131$) or bilingual exposure ($\chi^2(1) = .017, p = .897$). Independent samples t-tests confirmed that the two groups were equivalent in terms of mean parental education level ($t(35.226) = -.838, p = .408$) and age at assessment ($t(23.079) = -1.192, p = .245$).

Bivariate Comparison of Continuous Scores

Descriptive statistics for ADOS II Total Scores and subscale domain scores are shown in Table 14. Exploration of the data using histograms and normality statistics indicated notable positive skew across ADOS II summary scores,

particularly on the Social Affect scale ($z= 3.38, p<.001$). Application of a square root transformation (SQRT) produced an adequately normal distribution of ADOS II Total scores in both groups. Transformations did not improve normality of SA or RRB subscale distributions – and thus non-parametric tests were employed for between-group comparison of these subscale scores (Field, 2005). Examination of boxplots highlighted one ADOS Total score outlier in the VP group. In order to preserve power where possible and avoid potential exclusion of a legitimate VP population score, the bivariate analysis was run with and without this data point. Exclusion led to no change in the significance of findings, and as such reported findings are based on the full sample.

Table 14 ADOS Total and Subscale Scores at 30-34 months

		Very Preterm (N = 22)	Full-Term Control (N = 18)
Total Score	Mean	6.59	3.00
	SD	5.07	2.97
SA Domain	Mean	5.59	2.11
	SD	4.15	2.30
RRB Domain	Mean	1.41	.89
	SD	1.40	1.23

To explore the a-priori hypothesis that VP infants would show more traits associated with ASD at 30-34 months, an independent samples t-test was carried out comparing ADOS II Total scores (SQRT) across the groups. Non-

parametric Mann Whitney U tests were performed to compare sub-domain scores. VP infants scored significantly higher than full-term infants in terms of ADOS II Total score ($t(38) = 2.738, p=.009$) and on the Social Affect scale ($Mdn = 3$) ($U = 87.5, p=.002, r = .48$). VP infants and full-term infants received equivalent scores in the RRB domain ($Mdn = 1$) ($U = 148.5, p=.180, r = .22$).

Comparison of Diagnostic Cut-offs

Diagnostic categories according to DSM-V criteria were assigned to each infant taking part based upon their Total ADOS score. This was simplified to two categories: 'concern' (i.e. at or above ADOS II cut-off for ASD⁷) and 'no concern'. As Chi Square is unreliable where expected frequency in any cell is <5, Fisher's exact test⁸ was employed to compare the proportion of infants scoring above cut-off in the two groups. Using a two-sided test, there was no significant difference between the two groups in terms of diagnostic categorisations ($p=.427$).

Exploration of Longitudinal Associations Relating to ASD

Selection of Appropriate Statistical Methods

The second set of research questions related to longitudinal associations between early infancy measures (AOSI/IBQ-R-VSF/attention disengagement)

⁷ ADOS II cut-off score is variable dependent on the individuals level of spoken language (see ADOS II manual – Lord et al., 2012).

⁸ Fisher's exact test uses exact estimates of frequency, and is therefore robust to small expected values (Little, 2013).

and ASD symptomology at 30-34 months in the VP group. As fewer than 20 VP infants have to date completed 6/12 month assessments *as well as* ADOS II assessments at 30-34 months, a multiple regression approach (as planned) was considered not to be feasible. There are various “rules of thumb” regarding sample size requirements for multiple regression, but even the least conservative advise a minimum of ten participants *per* predictor variable (Tabachnick & Fidell, 2007; VanVoorhis & Morgan, 2007); samples smaller than this can lead to biased estimates of ‘true’ R (Maxwell, 2000).

Instead, a series of exploratory correlational analyses were carried out to explore longitudinal associations with ASD outcomes (ADOS II scores at 30-34 months) and motor, language and cognitive outcomes (Bayley III composite scores at approximately 24 months). To limit the number of tests, early infancy measures were only included if directly related to the research hypotheses *and* if significant group differences had been observed. *Attention Disengagement* was also included because of its theoretical significance, having been linked to ASD outcomes relatively consistently in high-risk sibling groups. As this was an exploratory analysis and no strong a-priori predictions regarding directionality of associations could be specified, no formal correction for multiple comparisons were applied and two-tailed significance were used. The results of this analysis are summarised as a correlation matrix in Table 15, and key associations are outlined.

At 6 months, higher AOSI scores were associated with more restricted and repetitive behaviours (ADOS II) at 30-34 months ($r = 0.515$, $n = 16$, $p =$

0.041), and poorer motor outcomes on the Bayley III at 2 years ($r = -.628$, $n = 18$, $p = 0.005$). By 12 months, AOSI scores showed a marginal positive association with Total ADOS II scores ($r = 0.518$, $n = 14$, $p = 0.058$), and higher scores appeared to be most strongly associated with difficulties in the social affect domain ($r = 0.635$, $n = 14$, $p = 0.015$).

In this small exploratory analysis, attention disengagement ability during the first year of life showed no significant associations with ASD or developmental outcomes. Parent reported negative affect at 12 months also appeared unrelated to ASD outcomes at 30-34 months. A *positive* association was observed between IBQ-R-VSF negative affect ratings and Bayley III composite language scores, surprisingly suggesting that higher levels of negative emotion were associated with better language outcomes at 2 years.

Table 15 Correlation matrix summarizing Pearson R correlations between key early infancy measures and ADOS II (30-34m) and Bayley III (24m) outcomes.

		30-34m ADOS II Social Affect	30-34m ADOS II RRB	30-34m ADOS II Total	Bayley III Cognitive (Composite)	Bayley III Motor (Composite)	Bayley III Language (Composite)
AOSI 6m Total	Pearson R	.427	.515*	.400	-.398	-.628**	-.131
	Sig. (2-tailed)	.099	.041	.125	.102	.005	.604
	N	16	16	16	18	18	18
AOSI 12m Total	Pearson R	.635*	.106	.518¹	-.005	.113	-.271
	Sig. (2-tailed)	.015	.717	.058	.983	.656	.278
	N	14	14	14	18	18	18
12m Negative Affect (IBQ-R-VSF)	Pearson R	-.200	.023	-.174	.176	.129	.425*
	Sig. (2-tailed)	.385	.922	.450	.399	.538	.034
	N	21	21	21	25	25	25
6m Attention Disengagement Effect	Pearson R	-.111	-.460	-.395	-.263	.133	-.022
	Sig. (2-tailed)	.694	.085	.145	.343	.638	.937
	N	15	15	15	15	15	15
12m Attention Disengagement Effect	Pearson R	.081	.333	.155	-.074	-.039	-.359
	Sig. (2-tailed)	.773	.226	.581	.785	.886	.172
	N	15	15	15	16	16	16

*Significant at the $p < .05$ level

** Significant at the $p < .01$ level

¹Marginally significant at the $p < .01$ level

Exploratory Partial Correlation

Three further partial correlation analyses were performed to explore the relationship between AOSI scores and ADOS II outcomes whilst controlling for level of cognitive, motor and language development. *N* values (and therefore *df*) were very small in these analysis – and thus findings should be interpreted with caution.

After controlling for cognitive, language or motor developmental level at 2 year follow-ups, all associations between 6 month AOSI scores and RRB domain outcomes in VP infants were no longer significant (all $p > .05$). In contrast, after partialling out developmental composite scores, associations between AOSI scores at 12 months and later emerging ASD symptomatology appeared to be sustained. After controlling for Bayley III language composite scores, higher scores on the AOSI at 12 months were still associated with greater social affective impairment at 30-34 months ($r = .644$, $n = 11$, $p = 0.044$), and the positive relationship with ADOS Total scores remained marginally significant ($r = .620$, $n = 11$, $p = .056$). Similarly, 12 month AOSI scores continued to be positively associated with ADOS II Social Affect ($r = .774$, $n = 11$, $p = 0.009$) and Total scores at 30-34 months ($r = .749$, $n = 11$, $p = 0.013$) after the influence of motor development was partialled out. Finally, controlling for cognitive level at 24 months actually appeared to strengthen associations between AOSI scores and both ADOS II social affect ($r = .819$, $n = 11$, $p = 0.004$) and Total scores ($r = .786$, $n = 11$, $p = .007$). This tentatively suggests that general cognitive abilities may have been *moderating* (i.e. limiting) the

relationship between VP infants' performance on the AOSI at 1 year, and later ASD outcomes.

Summary of Findings from Correlational Analyses

In summary, correlational analyses suggested that AOSI scores at 6 months and at 12 months were associated with ADOS outcomes at 30-34 months. Greater impairment on the AOSI in 6 month old VP infants appeared to be related to increased difficulties in the *non-social* ASD domain at 30-34 months. However, these associations no longer remained after developmental level was factored in. A particularly strong relationship was observed between observed impairments on the AOSI at 12 months, and social affective difficulties at 30-34 months; moreover, these associations appeared to persist after accounting for delays in motor, cognitive and language development.

Chapter 4: Discussion

The aim of the current study was to determine whether proposed early 'markers' of ASD symptomatology are identifiable in infants born VP, and consider how these markers relate to later ASD symptomatology in this group. This chapter will first summarise the key findings in relation to the main research hypotheses. Subsequently the results will be discussed in the context of the existing literature, addressing cross-sectional comparisons on the three infancy measures (the AOSI; attention disengagement task; and parent reported temperament) and the ADOS II at 30-34 months, as well as associations between infant markers and ASD outcomes. The following sections will consider strengths and limitations of the study, and outline possible avenues for future research. Theoretical and clinical implications of findings will be discussed.

Overview of Findings Relating to Study Hypotheses

A key research hypothesis was that infants born VP would show impairments relative to controls on a multiple marker measure of social and non-social behaviour (the Autism Observation Schedule for Infants [AOSI]; Bryson et al., 2008) at both 6 and 12 months. Study findings were in line with this hypothesis, with infants in the VP group receiving higher ratings overall and screening positive more frequently than full-term controls. As anticipated, it also appeared that these impairments were independent of group differences in general developmental level. In relation to longitudinal follow-ups, it was hypothesised that, in the VP group, impairments on the AOSI at 12 months

would be associated with higher levels of ASD related symptomatology on the ADOS II at 30-34 months. Preliminary findings suggested that this was the case; these associations appeared to persist after developmental level (motor, language and cognitive functioning) was taken into account.

In addition, infants born VP were expected to show greater difficulties with attention disengagement during the first year of life relative to controls, as well as atypical temperament profiles. The first of these hypotheses was *not* confirmed by the findings, with VP infants performing equivalently to full-term controls in terms of attention disengagement skill at both 6 and 12 months. However, some atypicality in early infant temperament profiles was observed; VP infants were reported by their parents to show higher levels of negative affect on the IBQ-R-VSF. No associations were observed, however, between either of these single marker measures and ASD outcomes at 30-34 months in the preterm group.

Finally, at 30-34 month follow-ups, VP infants were expected to present with relatively more symptoms associated with ASD compared to full-term controls. Using a semi-structured diagnostic assessment (ADOS II; Lord et al., 2012), VP infants did show marginally more symptoms associated with ASD at 30-34 months than control infants, with impairments most prominent in the domain of *social affect* (Lord et al., 2012). Elevated symptoms relating to restricted and repetitive behaviors were *not* observed in the VP group. Using ADOS II diagnostic cut-offs, there was also no evidence from this sample to suggest

that VP infants scored within the ASD clinical range more frequently than controls.

The following sections discuss these findings in relation to the research questions in finer detail, within context of the existing preterm and ASD infant-sibling literatures. Implications with regard to screening for ASD in preterm infant populations, and more widely in high-risk infant groups, are highlighted.

Autism Observation Scale for Infants (AOSI)

Cross Sectional Comparisons

As risk for ASD symptomatology is known to be significantly raised in VP/EP groups (Johnson & Marlow, 2011), it was anticipated that impairments would be observed on the AOSI during the first year of life. In line with expectations, VP infants showed quantitatively higher scores on the AOSI than full-term controls at both 6 and 12 months. Furthermore, using categorical cut-off values for elevated risk (≥ 7 risk 'markers') (Bryson et al., 2008), infants in the VP group also screened positive more frequently than full-term controls. These findings suggest that high-risk VP infants, like high-risk infant siblings, may show impairments on the AOSI long before the point at which an ASD diagnosis can typically be made (Yirmiya & Charman, 2010).

To our knowledge, this is the first time that elevated scores and screening rates on the AOSI have been demonstrated in a group of infants born in the higher risk, VP/EP range. Impairments relative to controls have, however, been reported previously in a lower risk preterm group; in a conference

submitted paper, Roncadin and colleagues (2010) reported that infants born moderately preterm (i.e. <37 weeks) received elevated scores on the scale relative to low risk, full-term controls. A recently published study also reported relatively high screening rates on the AOSI in a sample of moderate to high risk preterm infants, though the lack of a control group makes these findings difficult to interpret (Yaari et al. 2016). Their observed positive screening rates (21% at 8 months; 9% at 12 months) were noticeably lower than the present study, where around half of VP infants scored above cut-off for concern at 6 months and one third screened positive at 12 months. Perhaps the most likely explanation for this discrepancy is the lower average gestational age of the preterm population in the present study (26 weeks) relative to Yaari et al. (31 weeks). Research has consistently shown that the risk of neurodevelopmental impairment is closely linked to gestational age within the wider preterm population (Goldenberg et al., 2008). Moreover, it seems ASD symptoms specifically are more common in the lowest gestational age babies; Johnson & Marlow (2011), for example, found that gestational age of <25 weeks independently predicted ASD symptomology on the Social Communication Questionnaire (SCQ) in early childhood. The present findings suggest that very/extremely low gestational age babies may be particularly likely to score above cut-offs for concern on the AOSI assessment.

It is notable that observed positive screening rates on the AOSI in the present study were considerably higher than reported rates of *confirmed* ASD diagnosis in VP/EP children. In a whole population follow-up of infants born

extremely preterm, Johnson and colleagues estimated ASD rates of between 8% and 16% at age 8-11 years (2010b), which though still far in excess of incidence rates in the general population, is substantially below the positive screening rates of 30-50% observed on the AOSI in the VP/EP infants in this study. This is an indication that the scale may have relatively *poor* specificity in detecting ASD specific outcomes in high-risk preterm populations, and that false-positive screenings at 6 months may be unacceptably high. Moreover, it may suggest that the AOSI, like parental screening questionnaires (e.g. M-CHAT), struggles to differentiate early social communication difficulties from the broader ranging physical, cognitive and functional impairments that proliferate in this population (Hack et al., 2004; Kuban et al., 2009). However, it is notable that in the present study, birth status (VP/full-term) remained a significant predictor of AOSI score at 12 months after controlling for cognitive and motor functioning (Bayley III). This provides reassurance that the AOSI assessment *is* capturing some form of individual differences that are independent of general developmental level in young VP infants. To gauge whether these individual differences are ASD specific, it is therefore critical to consider associations with later emerging ASD symptomatology.

Associations between AOSI Scores and ADOS II Outcomes

It was hypothesised that VP infants' scores on the AOSI at 12 months, but not at 6 months, would be associated with ASD symptoms at 30-34 months. Correlational analyses indicated that these predictions were largely correct.

AOSI scores at 12 months showed a moderate⁹ correlation with ADOS Total scores at follow-up, and were even more strongly associated with impairments on the ‘social affect’ domain of the ADOS II. In line with predictions, 6 month AOSI scores were unrelated to overall ADOS II outcomes in the VP group. An unexpected association was found, however, between higher AOSI scores at 6 months and more ‘restricted and repetitive behaviours’ at 30-34 months. These preliminary findings suggest that, despite being developed and validated for use in a high-risk infant *sibling* population specifically, the AOSI assessment may be sensitive to specific emerging social communication difficulties in other high-risk groups.

To our knowledge, this is the first study to have considered associations between AOSI scores during first year of life and later ASD symptomology in a *high-risk* preterm group. However, its predictive capacities have been noted in a *low-risk* preterm sample. Roncadin and colleagues (2011) reported that the AOSI differentiated a group of low-moderate risk preterm infants ($N=49$) from full-term controls ($N=75$) at 12 months, and that inter-correlations with ADOS outcomes at 24 months were of a similar level to that seen in high-risk siblings. Though no study had yet extended these findings to a higher-risk preterm sample, Yaari and colleagues (2016) very recently reported that around 50% of moderate-high VP infants who show elevated scores on the AOSI at 12 months, then scored within the concern range on the ADOS

⁹ N.B. This correlation was ‘marginally’ significant.

Toddler module ('ADOS-T'; Luyster et al., 2009) at 18 months. However, the authors acknowledge that ADOS-T ratings at 18 months only indicate early ASD *risk*, as the stability of symptomatology at this relatively early age is unclear (Luyster et al., 2009). The present study extends these findings by showing that early impairments amongst VP infants on the AOSI are associated with ASD symptomatology later on in infancy, at an age when more stable diagnoses are typically feasible (Yirmiya & Charman, 2010).

The pattern of AOSI scores being associated with later ASD symptoms at 12, but not 6, months in this VP group is consistent with previous findings in high-risk infant sibling groups. Zwaigenbaum and colleagues (2005) found that total scores on the scale at 12 months predicted 24 month scores on the ADOS G in high-risk infant siblings (Lord et al., 2000), but similarly found that 6 month scores did not. More recently, Gammer and colleagues (2015) reported that AOSI scores at 14 months, but not at 7 months, were correlated with ADOS G scores in a high-risk infant sibling cohort (Lord et al., 2000). Together with these previous reports, the present findings indicate that the AOSI may have significantly more utility as an early screening measure for ASD by an infant's first birthday, than earlier. The fact that the scale has consistently failed to predict ASD outcomes in infants under 12 months old in different high-risk groups may suggest that items are not sufficiently specific to ASD outcomes at this young age. Moreover, the fact that all infants taking part (VP and full-term) tended to score higher at 6 months than 12 months is also indicative of this challenge with younger infants.

In the present VP cohort, higher scores on the AOSI at 12 months were more strongly associated with later emerging difficulties in the social affect domain than restricted and repetitive tendencies at 30-34 months. This pattern could be interpreted in a number of ways. Firstly, it may indicate that the AOSI is generally more sensitive to early behaviours associated with the non-social aspects of ASD. As previous studies have not differentiated between social and non-social symptom outcomes, this is difficult to evaluate relative to the existing AOSI literature. Equally, it could be argued that this discrepancy reflects a characteristic of VP infants as a cohort; the scale may be equally as sensitive to both non-social and social emerging symptoms, but VP infants primarily present with impairments in the social domain. It has previously been argued that social communication impairments in preterm groups may fundamentally differ in their nature and origin to 'idiopathic' ASD presentations (Johnson & Marlow, 2011) – and these findings may fit with these theories. However, it is also possible that the discrepancy simply reflects a difficulty in measurement of repetitive behaviours at 30-34 months either because these symptoms have yet to emerge, or because they are less easy to detect at this age (Cox et al., 1999). Further research with VP infants and other high-risk groups *including* follow-up at a later age would help to reach stronger conclusions about the specific profile of ASD symptoms the AOSI might predict in different infant groups.

Summary of Findings Relating to the AOSI

As expected from previous findings with lower risk preterm samples and infants at elevated genetic risk for ASD, infants born VP/EP showed impairments on the AOSI (Bryson et al., 2008) relative to control infants during the first year of life. Furthermore, it appears that high-rates of positive screening in the VP group were *not* simply due to elevated developmental delay. As expected, AOSI scores at 12 months, but not at 6 months, were associated with clinician rated ASD symptomology at 30-34 months. Together this suggests that the AOSI may have some utility as a screening measure in VP samples, as well as other high-risk infant groups, from around 12 months of age. As the stability and nature of ASD symptom profiles in VP groups is not well established, particularly during infancy, further research may explore associations between AOSI screening and ASD symptoms at an older age.

Attention Disengagement

It was hypothesised that VP infants would show impairments in attention disengagement relative to full-term controls at 6 and 12 months, and that impairments at 12 months would be associated with ASD outcomes at 30-34 months. Contrary to these hypotheses, no difference between the VP and full-term control infants was found in terms of disengagement of visual attention at either 6 or 12 months. There were also no associations between disengagement skill and ASD symptomatology at 30-34 months in the preterm group.

These findings add to a somewhat inconsistent literature regarding visual attention skills in young preterm groups. Though atypicalities in visual attention have been reported during the first year of life (e.g. Hunnius & Geuze, 2004), including longer look durations, slower disengagement, and slower attention shifts, other studies have reported typical disengagement or even *advantages* relative to full-term controls in lower risk infants (Hitzert, Van Braeckel, Bos, Hunnius, & Geuze, 2014; van de Weijer-Bergsma, Wijnroks, & Jongmans, 2008). In relation to the surprising observation of early advantages, Hunnius and colleagues highlight that the additional visual exposure preterm infants have gained compared to their full term peers may have accelerated the maturation of cortical processes involved in disengagement. The present findings, of difficulties with reflexive orienting (i.e. reaction times on Baseline trials) at 6 months, but no impairment in disengagement ability (a higher level cortical process), are perhaps consistent with this conjecture. Hitzert and colleagues (2014) also considered whether relative impairments in preterm infants may be mediated by perinatal complications and brain damage, which may add to inconsistency in group-level findings. Further research may therefore consider how individual differences amongst VP infants, relating to specific medical and neurological complications, are associated with visual disengagement difficulties.

The present findings contrast with findings from the ASD infant sibling literature, where impairments in disengagement of attention have consistently been observed relative to low-risk controls towards the end of the first year of

life, and associations with emerging ASD symptomatology have been documented (Bedford et al., 2014; Elsabbagh et al., 2009). ASD infant sibling studies have in particular highlighted the importance of a pattern of deterioration in attention disengagement skill (i.e. an increased tendency to get 'stuck' on visual stimuli) over the second six months of life in predicting ASD emergence (Elsabbagh et al., 2013; Zwaigenbaum et al., 2005). In the present study there was no evidence of any such regression in disengagement skills amongst VP infants, at the group level. Overall, this lack of group differences and association with ASD outcomes suggests that attention disengagement may *not* be a generalisable marker for emerging ASD symptomatology in preterm groups.

A broader implication of these results, and the contrast with findings in ASD infant sibling groups, is that early markers for ASD may not be homogeneous across different infant populations. As noted previously, research into the early trajectories towards ASD has almost entirely been informed by prospective studies with infant-siblings, and concerns have already been raised about the implications of this in terms of generalisability (Jones et al., 2014). The present findings add weight to these concerns. Given the vast heterogeneity known to exist in ASD presentations (Frith & Happé, 2005), it is perhaps unsurprising that infants show diverse developmental 'pathways' towards ASD symptomatology. The extent to which trajectories can be confidently predicted on the basis of specific risk-factors (e.g. prematurity vs. familial risk) is unclear without more extensive research with infants' from

different high-risk groups. Nonetheless, the current findings indicate that there may be some significant difference in trajectories between infants at familial risk and preterm groups.

Summary of Findings Relating to Attention Disengagement

In summary, research hypotheses relating to attention disengagement impairments in VP infants were *not* confirmed in this study. VP infants performed equivalently to their full-term counterparts at 6 and 12 months, and no associations were observed with ASD outcomes during the third year of life. Importantly, this suggests that attention disengagement may not be a universal early infancy marker for ASD symptomology. When contrasted with findings in ASD infant sibling populations, the current findings highlight that heterogeneity may be an important feature of developmental trajectories towards ASD symptomatology.

Infant Behaviour Questionnaire (Temperament)

A-priori research hypotheses stated that VP infants would present with atypical temperament profiles relative to full-term controls at 6 and 12 months, and that these atypical profiles may be associated with measurable ASD symptomatology at 30-34 months. The first hypothesis, regarding cross-sectional comparisons, was partially confirmed. Using the IBQ-R-VSF (Gartstein & Rothbart, 2003; Putnam et al., 2014), parents of VP infants reported higher levels of 'negative affect' compared to full-term controls at 12 months, but not at 6 months. Interestingly the significance of this group difference increased notably after controlling for gender in the analysis. This

observation appeared to reflect relatively *lower* levels of parent-reported negative affect amongst male infants – a finding that has been reported previously on such measures (e.g. Gartstein & Rothbart, 2003; Martin, Wisenbaker, Baker, & Huttunen, 1997). No difference were reported between the groups in terms of surgency (i.e. positive affect) or effortful control. The second hypothesis, relating to longitudinal correlations, was not confirmed by the data; there was no association between early atypicalities in infant temperament (negative affect) and ASD symptomatology at 30-34 months in the VP group. Together, this suggests that atypical infant temperament may not be an effective marker for emerging ASD symptomatology in VP groups.

The observation of subtle group differences in temperament profile at 12 months is somewhat consistent with previous research with preterm infants (Langerock et al., 2013) and high-risk infant siblings (Clifford et al., 2013).

Using the IBQ-R (Gartstein & Rothbart, 2003), Clifford and colleagues reported reduced positive affect and self-regulatory ability ('effortful control') in high-risk ASD siblings compared to controls during the first two years of life. In high-risk *preterm* groups, atypicalities have particularly been noted in relation to self-regulation and negative reactivity (Hughes, Shults, Mcgrath, & Medoff-Cooper, 2002; Langerock et al., 2013). For instance, Langerock and colleagues (2013) found that VP infants showed elevated reactivity in anger-eliciting situations, as well as *reduced* reactivity toward fear-eliciting situations in a clinician-rated behavioural observation paradigm at 12 months.

Interestingly these researchers found no group differences on the IBQ-R-VSF

at 12 months; but it is quite possible that this related to limited power, with only 21 full-term and 16 VP infants included in analyses (compared to 39 full-term and 35 VP in the present study). This might suggest that atypicalities in temperament, particularly relating to negative affect, are observable during the first year of life in VP samples, but these differences may not be easily captured using a parent-report approach.

The observed finding of *no* association between early parent-reported temperament and ASD outcomes in the present study contrasts with reports in the infant sibling literature using similar methods. In prospective research with high-risk infant siblings, associations between atypical early temperament profiles and ASD-related outcomes have consistently been reported (Clifford et al., 2013; Garon et al., 2009; Zwaigenbaum et al., 2007). Clifford and colleagues, for example, reported a pattern of increased perceptual sensitivity, increased negative affect and reduced tolerance of physical affection ('cuddliness') in infant siblings, which predicted later emerging ASD symptoms. Similarly, Zwaigenbaum and colleagues (2005) found that passivity and decreased activity levels at 6 months, and extreme distress reactions and reduced positive affect at 12 months, were predictive of later ASD symptomatology. In the context of the wider literature, the present findings therefore suggest that parent-reported infant temperament may not generalize from infant sibling populations as an effective early ASD marker. This again highlights that extensive heterogeneity might exist in

developmental trajectories towards ASD symptomatology, and that these trajectories may differ between infants at familial risk and preterm groups.

It is possible that the use of the very short form of the IBQ-R in the present study, instead of the full form IBQ-R, played a role in the lack of longitudinal associations with ASD symptomatology. The IBQ-R-VSF was employed to minimise the research burden on parents, especially given they were required to complete numerous other questionnaires as part of the wider research programme. Compared to the full-scale Infant Behaviour Questionnaire-Revised (191 items) (Gartstein & Rothbart, 2003), the very short form has substantially fewer items (only 37), and a much simpler scale structure derived through factor analysis of the 14-scale IBQ-R (Putnam et al., 2014). Psychometric analysis has, however, demonstrated that the IBQ-R-VSF is strongly correlated with the long-form version, and has only slightly lower levels of internal consistency and longitudinal stability (Putnam et al., 2014). As a result, findings in the present study should be broadly comparable with infant sibling studies that have employed longer versions of the scale.

Though no associations with ASD outcomes were observed, exploratory analyses did highlight a significant association between negative affect scores on the IBQ-R-VSF at 12 months, and *better* Bayley III language scores at 30-34 months. In the context of the broader literature on infant development, this association is rather surprising, as negative emotionality during infancy has been shown to predict less favourable language outcomes (Bloom & Capatides, 1987; Salley & Dixon Jr, 2007; Dixon Jr & Smith, 2000; Noel,

Peterson, & Jesso, 2008). Therefore, especially given this difference was only marginally significant, it is quite possible that the finding represents a Type 1 error. Future research may wish to explore the relationship between early temperament and subsequent language development in VP populations further, however, as it is conceivable that distinctive trajectories exist.

Conceptual and Methodological Challenges Associated with 'Infant Temperament'

Various authors have called into question the construct validity of infant temperament, in terms of what it is *actually* measuring (Henderson & Wachs, 2007). The constructs assigned to 'temperament' in the literature range from quite circumscribed to much more broad according to different theoretical perspectives (Goldsmith et al., 1987). Whereas some argue that it is best conceived as a set or profile of specific behavioral dimensions (Rothbart, Derryberry, & Posner, 1994), it has also been conceptualised as something closer to a 'personality' attribute that distinguishes different 'types' of individuals (Kagan, Snidman, Arcus, & Reznick, 1994). The IBQ-R-VSF aligns with the 'behavioural profile' approach, which although more biologically based, could still be argued to lack specificity in terms of the functions it measures – especially when compared to a functionally and neurally more well-defined process like attention disengagement (Posner, 2014). So whilst there is evidence that temperament questionnaires measure individual differences that are somewhat stable over time (Casalin, Luyten, Vliegen, & Meurs, 2012) and which show association with developmental and psychiatric outcomes (Sayal, Heron, Maughan, Rowe, & Ramchandani, 2014), it is

possible that conceptual difficulties limit the generalizability of ‘temperament’ as a predictor of highly *specific* infant outcomes across diverse infant groups.

Reliance on parent report in this study, and in the vast majority of research into infant temperament, also represents a challenge to interpreting findings. Questionnaire approaches are affected by various factors that may impact reliability in reporting – including understanding/interpretations of questions, memory and knowledge of behaviours, personal response sets and social desirability, as well as differences in explicit and implicit reference points for judging infant behaviours (Rothbart & Mauro, 1990). Reports of poor consistency between temperament questionnaires and direct observations of infant behaviour (e.g. Langerock, 2013; Seifer, Sameroff, Barrett, & Krafchuk, 1994) highlight that these factors may have a significant impact upon the reliability of parent reports. These sources of bias become even more important when comparing temperament across discrete infant populations, such as preterm and full-term infants, where parents may have *systematically* different expectations of their infant’s behavioural profile or temperament. For example, it is possible that parents of VP infants would rate their infants behaviour relatively more positively, in the context of complex (and possibly ongoing) medical or developmental difficulties. Similarly, systematic biases might be anticipated in ASD infant siblings populations, by virtue of already having raised an infant with ASD. Though the impact of these biases is difficult to quantify, they should be recognised as a potential challenge to parent-reported infant temperament as an early ‘marker’ for ASD.

Summary of Findings Relating to Infant Temperament

In summary, the current study found higher levels of parent-reported negative affect, but no other differences in early temperament, in VP infants relative to their full-term peers. Unlike relatively consistent findings with high-risk ASD infant siblings, there was no evidence of association between early temperament and ASD symptomatology in the third year of life. However, there were a number of limitations that must be taken into account when drawing conclusions from the findings - most notably, the reliance on a relatively brief parent-report questionnaire of temperament, which may be subject to systematic response biases. In the context of the wider literature, these findings suggest that parent-report measures of infant temperament may *not* generalise beyond ASD infant sibling groups as a marker for emerging ASD – and, specifically, not predict later ASD symptomatology in VP populations. This again highlights the probable heterogeneity in early markers for ASD across different high-risk infant groups.

ASD Symptomology at 30-34 months (ADOS II)

The final study hypothesis related to social communication outcomes at 30-34 months, with infants in the VP group expected to present with more symptoms associated with ASD than full term controls. In line with this hypothesis, infants in the VP group did score quantitatively higher on the ADOS II (Lord et al., 2012) than full-term controls at the final follow-up; this difference was tentatively observed in terms of Total score, and more clearly observed on the Social Affect subscale of the ADOS. Interestingly, VP infants did not score

higher than the controls in the restricted and repetitive behaviour domain. There were also no overall group differences in terms of the absolute number of infants falling above ADOS II cut-off criteria for Autism Spectrum Disorder. It is important to recognise that the limited sample size at follow-up may have influenced the pattern of observed group differences, however – particularly where null or marginal findings were obtained.

A quantitative elevation in ASD symptomology was anticipated in the VP group on the basis of previous studies with preterm samples using the ADOS-G (i.e. the predecessor of the ADOS II) (De Groot et al., 2006; Pritchard et al., 2016) and the ADOS Toddler Module (Yaari et al., 2016). This finding is also consistent with outcomes of larger scale epidemiological studies that have shown a quantitative shift (i.e. increase) in parent reported ASD symptom presence using parent report instruments such as the Social Communication Questionnaire (SCQ; Chandler et al., 2007) (S. Johnson et al., 2011). The present findings confirm that this quantitative elevation in symptomatology is also observable using a standardized clinical behavioural observation approach at a relatively early age.

ASD Diagnostic Categorisations

In the present study, around 20% of the VP infants scored at or above ‘cut-off’ for ASD on the ADOS II, which is notably higher than estimated rates of confirmed diagnosis in VP/EP populations. In their whole population follow-up, Johnson and Marlow estimated that between 8 and 16% of EP children at age 11 meet diagnostic criteria for ASD (2010a). Another report published very

recently from a multi-centre follow up study of over 800 children born at <28 weeks in Australia (the “ELGAN” study) estimated confirmed ASD diagnoses of approximately 7% (Kuban et al., 2016). Three possible explanations for this discrepancy are discussed.

It is possible that the higher estimate of ASD in the present study reflected genuinely high rates of ASD in this specific VP cohort. It is of note that all VP infants were recruited from a tertiary neonatal intensive care unit, which is indicative of a medically high risk group. Evidence has increasingly demonstrated that neonatal morbidity is a stronger predictor of a wide range of adverse outcomes than gestational age itself (Schmidt, Asztalos, & Roberts, 2003). As such, the inclusion of only VP/EP infants who required prolonged specialist medical care may have led to particularly high detection rates for ASD. It would be interesting to consider how early indices of morbidity (e.g. number of days spent in specialist inpatient care) in this cohort related to individuals ASD outcomes. The other key feature of the cohort that may have inflated ASD estimates was the over-representation of boys in the sample (15 of the 21 VP infants who completed the ADOS II were male). It is widely recognised that ASD rates are higher in males than females, with an estimated population wide ratio of 4:1 (Fombonne, 2003). Though this ratio is thought to be lower in EP populations (S. Johnson et al., 2010a), perhaps reflecting unique risk profiles in this group, boys born EP are still approximately twice as likely to be diagnosed with ASD than girls (Kuban et

al., 2016). The unexpectedly high rates of infants above diagnostic cut-off in these assessments may therefore reflect site and sample specific factors.

It is also possible that the relatively early age of follow-up inflated positive diagnostic outcomes on the ADOS II. Though there is evidence that reliable ASD diagnoses can be, and frequently are, made from age two (Cox et al., 1999; Stone et al., 1999), there is consensus that the stability of diagnosis and symptoms generally improves with age (Moore & Goodson, 2003). As such it is quite possible that some VP infants who scored above cut-off at 30-34 months would no longer do so later in childhood. Though research has shown that instability in diagnostic categorisation is more often due to early *under-detection* rather than over-detection in the general population (Lord et al., 2006), it is not clear whether this generalizes to globally at-risk groups. The ADOS II has not been validated specifically for use in VP populations, and it may be that there are problems of specificity and 'over detection' in younger groups. The reliability and stability of early diagnostic categorisations in VP/EP populations at this young age warrants further investigation.

The final factor that may have influenced the unexpectedly higher estimates of ASD is the assessment method used (i.e. behavioural observation only). The ADOS II is considered to be one of the "gold standard" assessment tools in ASD, and is widely used by both clinicians and clinical researchers (Hurwitz & Yirmiya, 2014). It is often favoured in research contexts as it provides a standardised estimate of symptom severity in line with DSM criteria, is not reliant on parent report, and is administered by trained practitioners (Lord et

al., 2012). The ADOS II is not, however, a *complete* diagnostic assessment, and would *not* be used in isolation for clinical diagnostic purposes because it only provides a ‘snapshot’ of an individual’s behavior. It should therefore be used only very cautiously as an estimate of diagnostic categorizations. NICE (2011) recommends that a thorough *diagnostic* assessment for ASD should include both a structured parental interview including a detailed history taking (e.g. the Autism Diagnostic Interview; ADI/ADI-R) as well as direct clinical observation of behavior¹⁰. This highlights that estimations of diagnostic category in the present study should be interpreted with caution. Future research should, where possible, include both parental interview and direct observation in order to provide more clinically valid and reliable estimates of ASD diagnosis.

DSM V: Discriminating Social Affect and Repetitive Behaviours

The present study was one of the first to measure ASD symptomology in a preterm sample using the ADOS II (Lord et al., 2012), as opposed to parent questionnaire or screening methods or the previous version of the ADOS. This permitted an exploratory examination of the symptom profiles in the VP group according to *current* conceptualizations of ASD according to DSM V (American Psychiatric Association, 2013). DSM V introduced some significant

¹⁰ For this reason, ADOS scores were *not* used for *clinical* diagnostic purposes in the present study and this was made clear to parents taking part. In line with ethics approved procedures, any child that did score above cut-off for ASD on the ADOS was notified to the study Principal Investigator who contacted families to offer an onward referral for further assessment.

changes to the conceptualisation and measurement of ASD (Volkmar & McPartland, 2014) including the collapse of multiple diagnoses (Autism, ASD, Asperger's, PDD-NOS) into one single diagnostic category ("Autism Spectrum Disorder"), and a shift from recognizing three separate domains of impairment in the ASD profile to just two: 'social communication and interaction' and 'restricted, repetitive patterns of behaviour, interests or activities'. In order to reach diagnostic criteria for ASD diagnosis, individuals must now show difficulties in *both* social and non-social domains. Using the ADOS II thus made it possible to consider whether VP infants, at this relatively early age, present with this 'complete' profile. The finding of group-level impairments relative to controls on the social affect scale, but not in the restricted and repetitive domain, may suggest that VP infants actually present with just 'one half' of this full behavioral profile. This is important because children without any repetitive or restricted behavioural tendencies would no longer reach diagnostic criteria for ASD under DSM V (American Psychiatric Association, 2013).

The observed profile is, however, more likely to be a reflection of methodological factors. As noted previously, it has been reported that aspects of the 'restricted and repetitive behaviour' domain are less reliably detected at earlier ages (Lord, Storoschuk, Rutter, & Pickles, 1993). Research has shown that whilst repetitive behaviors are usually reliably detected by 3-4 years old, they are rarely observed in two year olds (Cox et al., 1999) - with social deficits and delays in spoken language generally thought to be the most

prominent difficulties in very young children with autism (Stone et al., 1999). In this sense, the observed pattern of ASD symptoms in the VP group is quite typical of early ASD profiles. Another methodological factor that may have contributed towards the observed profile was the use of the ADOS, which can struggle to fully capture repetitive and restricted tendencies within only a brief sample of behaviour (Lord et al., 2006). As such the most likely explanation for the observed 'profile' of impairment in the VP group is the early age of follow-up and reliance on behavioural observation. Further follow-up at a later age, using both behavioural observation *and* a diagnostic interview schedule to gather historical and current diagnostic information from parents, would help to test this hypothesis.

Summary of Findings on the ADOS III

In summary, the study findings indicated that at 30-34 months, infants born VP show quantitatively more symptoms associated with ASD than full-term controls, using a gold standard clinical behavioural observation tool. Using this approach, somewhat higher than expected rates of 'ASD' were detected according to ADOS II cut-offs – which may relate to cohort effects, the relatively early age of follow-up and/or reliance on behavioural observation only for estimation of diagnostic outcomes. VP infants appeared to present with greater impairments in social communication skills at 30-34 months than in the domain of repetitive tendencies. Further research using a more thorough assessment approach which includes a diagnostic parent interview

and behavioural observation in older VP children would help to clarify the longitudinal stability of this symptom profile.

Strengths and Limitations

Strengths of the Study

One of the key strengths of the present study was its attempt to bridge two quite disparate literatures – that of early markers and diagnosis in autism spectrum disorders and outcomes after preterm birth. There have been significant advances in our understanding of early markers for ASD in recent years, which has resulted in the development of promising tools for behavioural observation (Yirmiya & Charman, 2010). However, the reliance almost entirely on research programmes involving one high-risk infant population (ASD siblings) has meant that the generalisability of these infant measures and markers is relatively unknown (Jones et al., 2014). By recognising recent research that has highlighted VP birth as a significant risk-factor for ASD, and applying the ‘infant sibling’ prospective research design, it has been possible to start gaining insight into this generalisability. As such, the bridging of two literatures has added to understanding of both outcomes after preterm birth, and early trajectories associated with ASD, as well as how these interrelate.

There were several aspects of this study that were novel. First, to our knowledge, it was the first study to measure the performance of a *high risk* preterm (i.e. VP/EP) infant sample on both the AOSI (Bryson et al., 2008) and the ADOS II (Lord et al., 2012). However, more importantly, this was the first

attempt to consider relationships between preterm infants' performance on a range of infant markers (including the AOSI) and ASD related outcomes according to DSM-V criteria in a VP group. Though the scope of these investigations were curbed by sample sizes in the longitudinal follow-up, the correlational analyses provided a promising indication of measures that might be more effective at detecting early ASD related symptomatology in VP groups, and that warrant further investigation.

More generally, it could be argued that the longitudinal research design was a study strength. The advantages (and disadvantages) of longitudinal data have been debated at length, but some key benefits are it's ability to establish temporal order, measure change and make stronger hypotheses about causal influences (Rajulton, 2001). In this study, collecting multiple measurements on related measures at different time points in the same group of infants, enabled some consideration of development (i.e. change) over time. Though only preliminary considerations about associations between infant measures and outcomes at 30-34 months were possible with the limited number of infants who have been followed up to date at 30-34 months, a cross-sectional design would have precluded *any* analysis of relationships over time.

It is notable that retention rates between 12 and 30-34 months, thus far, have been reasonable (82% in VP group). This is especially true given the considerable time-cost to families of taking part and the many potential obstacles to retention in this cohort (e.g. parents returning to work; subsequent pregnancies; need for child-care; moving home). Various factors

are likely to have contributed towards effective retention, including strategies to help create a sense of community and continuity for parents (e.g. a study website) and the active collection of feedback from parents on their experience of taking part which was then used to inform subsequent follow-up planning. Retention was evidently better in the preterm group, and informal discussions with parents indicated that this might have reflected their vested interests and positive experiences of care within the service, which resulted in greater motivation and willingness to 'give back' to clinical research.

Limitations of the study

Sample Sizes

The most notable limitation of the study was the lower than anticipated sample size included in longitudinal follow-ups at 30-34 months. This was primarily due to practical circumstances associated with the wider research programme, which led to follow-up being moved from 24 to 30-34 months. As such, fewer infants than expected reached follow-up age during the time-scale of the DClinPsy project; whereas a sample of approximately 30 infants per group were anticipated in both groups at 24 months, only 22 preterm and 15 full-term control infants reached 30 month age criteria by mid-April 2016. It is notable that as data collection is ongoing, the remaining infants will still be followed-up over the summer and autumn. The primary impact of the limited longitudinal follow-up sample was upon the selection of statistical procedures that could be employed to explore associations between infancy markers, ASD symptomatology and developmental delay. A hierarchical regression

approach, using theoretically indicated predictor variables, would have enabled a more detailed consideration of how infant marker variables and developmental level independently *and* cumulatively contribute towards VP infants' performance on the ADOS. Without these analyses, it is difficult to draw any strong conclusions about the role of developmental impairments in early ASD outcomes amongst VP cohorts. However, correlational analyses have provided an informative indication of which infant measures may be the most promising as 'markers' of ASD symptomatology in VP samples; findings suggest that the AOSI may be particularly sensitive to emerging difficulties with social communication. Future research (including continued follow-up with this study cohort) is encouraged to investigate this relationship further.

There were also some advantages to following up at 30-34 months, rather than 24 months. The most notable benefit is that the reliability of ASD assessment measures (including the ADOS) is known to improve significantly through the third year of life (Charman & Baird, 2002). As such, the estimated ASD symptom scores (and diagnostic categorisations) obtained at 30-34 months are *more* likely to be stable and reliable indications of outcome than at just two years – even if impairments in the repetitive domain are still difficult to capture (Cox et al., 1999).

There was also a significant amount of assessment data missing due to incomplete, unusable or missed assessments, which affected sample sizes throughout cross-sectional and longitudinal analyses. This problem largely applied to data collected directly from infants, with substantially more

comprehensive data-sets achieved on the parent-reported IBQ-R-VSF, compared to the attention disengagement task and the AOSI. The challenges of obtaining good quality data from young infant participants are widely recognised, even where extensive adaptations to research procedures and settings are made to account for limitations in infants physical, cognitive, and emotional development (Field & Behrman, 2004). It could be argued that these challenges are amplified when working with high-risk groups who are likely to have additional sensory, regulatory and physical difficulties. Nonetheless, the proportion of missing or unusable assessment data was still greater than anticipated prior to data audit and cleaning.

Though it is unclear exactly why missing data was such a problem, it can be speculated that the complexity and scale of overarching study procedures associated with the wider research programme played a role, as they placed large demands on infant testing sessions. At each research visit, the wider programme involves data collection on multiple assessments (i.e. not only those included in this study) across a number of hours. Unfortunately, it seems that this breadth of data collection may come at the cost of a more comprehensive data-set on each *individual* assessment. Future large-scale research programmes of this sort may therefore consider fine-tuning protocols to a smaller number of assessments, or alternatively, finding ways of building in more flexibility in testing (e.g. allowing for multiple visits). Obviously, the latter would involve significant cost and time implications for both researchers and parents, and the pros and cons of this would have to be carefully

considered. However, it is clear that future research should prioritise strategies to promote infants' level of engagement and willingness to cooperate in order to maximise the quantity and quality of data obtained.

Group Matching (SES & Gender)

Another limitation of the study was inadequate matching of the VP and full-term groups on demographic indices at 6 and 12 months. The two groups showed significant differences in terms of both gender distributions and estimated social economic status – and it should be recognised that either (or both) of these discrepancies may have affected findings, particularly with regard to cross-sectional comparisons. The greater proportion of males in the VP group compared to the full-term group was particularly problematic, given the known male bias in ASD rates in the general population (approximately 4:1) (Baird et al., 2006; Mandy et al., 2012). Though male:female ratios for ASD in preterm populations appear to be somewhat lower than in the general population, gender biases in diagnosis rates do still exist, with approximately twice as many boys reaching diagnostic criteria than girls (S. Johnson & Marlow, 2011). However, as the gender imbalance in the VP group was likely to have arisen by chance (see previous chapter, p. 72-73), it was justifiable to carry out statistical analyses to 'control for' this problematic group difference (Miller & Chapman, 2001). These analyses provided reassurance that gender *probably* did not play a significant role in group differences on the AOSI and IBQ-R-VSF. Nonetheless, reliance on statistical methods to control for possible confounds does limit the certainty with which conclusions can be

drawn. As Miller and Chapman (2001) point out, statistical control for potential confounds cannot replace careful sampling and avoidance of non-matched groups in the first place.

Systematic group differences in estimated parental education level also represent a barrier to interpretation, but it is likely that this group difference is genuinely reflective of the clinical population. Parent education level is generally considered to be a relatively good proxy for SES (Oakes, 2006), and it is well established that socioeconomic disadvantage is associated with an increased risk for premature birth (Behrman & Butler, 2007). As such, prematurity and social economic disadvantage are inextricably linked, and controlling for this group difference via ANCOVA would be both statistically and theoretically unsound (Miller & Chapman, 2001). Though efforts could have been made to match the two groups on SES (by specifically targeting a lower SES control group), this may have led to underestimation of impairments. Comparing a clinical sample with a 'typical' non-clinical sample (as opposed to an SES matched sample) is arguably preferable, as it fully captures the real-life variance associated with prematurity. Nonetheless, the lack of correlation between parent education level and AOSI scores or IBQ-R-VSF *Negative Affect* ratings in this sample does suggest that SES is unlikely to have played a significant role in these group level findings.

Together these limitations highlight the complexities in interpreting findings from clinical comparison groups. Future research would benefit from maximizing equivalence between groups, particularly on demographic indices

that are *not* intrinsically linked to that clinical population – such as gender, in this case. It is recognized, however, that this can be difficult to achieve when recruiting from relatively rare clinical samples; in the present study, greater selectivity in recruitment from the test site would have come at the expense of valuable participant numbers.

ASD Outcome Measurement

The use of the ADOS II, a “gold standard” in diagnostic assessment of ASD symptomatology, could be considered a strength of this study (Lord et al., 2006, 2012), particularly in light of the novel application to a high-risk VP population. However, the use of this scale *in isolation* could also be considered a limitation. The most notable drawback of this design is that direct behavioural observation, even when scored by multiple trained raters, is not sufficient to make valid clinical judgements regarding diagnosis. Clearly, though the ADOS provides a rich sample of clinically relevant information, it can only provide a snap-shot of an individual’s behaviour. As such, in a clinical context it would always be used in combination with a parental interview including a detailed history taking (Lord et al., 2012; NICE, 2011; Pilling et al., 2012).

Though practical and time restraints made it impossible to include a parent interview in the present study, a richer and more reliable picture of outcome would have been achieved if this information had been available. The Autism Diagnostic Interview-Revised (ADI-R) (Lord et al., 1993) is widely considered to be gold standard tool for parent interview. However, future research with

high risk preterm groups may wish to consider the computer-based “3Di” as an alternative tool (Skuse, 2004), as it takes into account and screens for a range of possible comorbidities. In a VP sample, who are known to be at risk for a wide range of developmental and psychiatric outcomes, this additional information could be highly valuable.

It is important to note that all clinical scales used within this study – including the IBQ-R-VSF, the AOSI and the ADOS II – have not been specifically validated for use within VP populations. Therefore, an overarching caution should be used when interpreting the present results, and more broadly when reflecting on findings from these scales in high-risk preterm samples. Clearly, this reflects broader difficulties associated with accurately measuring specific outcomes in globally at-risk samples. However, future research is warranted to provide validation and, if necessary, modification to these assessment instruments for use in VP populations.

Theoretical and Clinical Implications

The current findings add to the growing body of evidence suggesting that VP infants are at risk for presenting with symptoms associated with autism spectrum disorders, at least in the social affective domain. Moreover, though further research with larger sample sizes and longer term follow up is clearly required, these preliminary findings suggest that there may be clinically feasible methods of identifying those infants at highest risk earlier in infancy. At the least, it would appear that there is some consistency in those infants who show early difficulties on the AOSI from around 12 months and those

who present with impairments in social affective functioning during the third year of life. The potential benefits of early identification and intervention in terms of mood, functioning, family stress and even symptom severity outcomes in ASD are increasingly highlighted (e.g. Estes et al., 2015). Therefore, detecting these more specific difficulties earlier in preterm groups could be beneficial to both children and families, in order to support families in accessing appropriate support.

On a practical level, these findings highlight that clinicians working with children and young people should be aware of the possibility of social communication difficulties in preterm children, particularly in those born very early. This becomes particularly pertinent when we consider that more and more children are surviving VP/EP birth, and that it is these infants at the limits of viability (i.e. 22-24 weeks) that have the poorest prognoses on a wide range of developmental and specific outcomes (Saigal & Doyle, 2008).

However, it is also important for clinicians to recognise how these symptoms or difficulties fit into the broader 'preterm behavioural phenotype' (S. Johnson & Marlow, 2011), and thus how wider problems may complicate typical assessment approaches and make specific impairments less easy to detect.

At the same time, it is important to remember that these infants often have complex medical and developmental histories and the potential harms of 'over' labelling with additional 'disorders' or difficulties may be reflected upon.

At present, child development services in the UK do not carry out formal routine screening for social communication difficulties specifically in VP/EP

populations, though infants are typically offered broader development screening (i.e. Bayley III or similar) through to two years (NICE, 2015). The present findings highlight that relying on broad developmental follow-ups only may 'miss' some more specific social communication difficulties which are independent of motor, cognitive and language delays. Therefore professionals following the development of preterm infants may consider assessing social-communication and related behaviors in routine follow-up of preterm infants; and the present findings indicate that clinician led multiple marker behavioural assessments like the AOSI may be the most viable avenues to explore.

Further research, however, is required into the best approach to screening. It is widely documented that poor specificity and false-positives are a problem when using parent-report screening methods (e.g. the M-CHAT) in VP/EP populations due to wider motor, sensory and cognitive impairments (Kuban et al., 2009; Luyster et al., 2011); unfortunately, the high rates of positive screening at 6 and 12 months in the present study suggest that the same may be true with the AOSI. Indeed, it may simply be that new or adapted tools need to be developed and validated in preterm samples, which can effectively isolate ASD specific impairments from the the wider difficulties these children have. Nonetheless it could be argued that a direct observational measure like the AOSI is preferable to parent report, as trained clinicians are likely to be more skilled at differentiating ASD specific markers that may be qualitatively different to more 'typical' motor control, sensory, postural and regulatory difficulties seen preterm populations.

Finally, there are some broader clinical implications of these findings, relating to wider efforts to understand developmental pathways towards ASD. Two infant measures that have been shown to predict ASD outcomes in ASD infant siblings – attention disengagement skills and the IBQ-R-VSF - showed no evidence of association with these same outcomes in VP infants. Though it is important to recognise that methodological discrepancies may have played a role in this pattern of findings (e.g. use of a short form of the IBQ, use of behavioural observation only for ASD assessment, and limited sample size in the present study), the current findings add weight to arguments that the ‘prodromal’ markers of “ASD” may not be uniform across all infants (Jones et al., 2014). That is, different risk factors for ASD may be associated with different, and perhaps specific, developmental trajectories. This heterogeneity is likely to be a significant challenge to efforts to improve early identification – and further research with different high-risk groups is required to help elucidate specific trajectories associated with specific risk factors.

Recommendations for Further Research

One of the most obvious ways in which the findings from this study could be extended is by carrying out further follow-up with VP/EP infants at an older age. The main benefit of this is that it would enable the longer-term stability of ASD symptoms observed in VP infants at 30-34 months to be explored. Given the discussed limitations of relying on ADOS II scores only, the use of both behavioural observation and a parental diagnostic interview would be advised,

in order to more accurately estimate diagnostic outcomes and long term symptom profiles (i.e. social affect/repetitive behaviours) (NICE, 2011).

Another avenue of further research that has been discussed is the consideration of specific profiles across individual items of the AOSI, and examining their relationship with ASD outcomes. Given overall AOSI scores *were* associated with ADOS II scores in this study, and this association appeared to be independent of general developmental delay, it would be interesting to explore which items appear to be driving associations with ASD specific outcomes. Though these analyses may primarily be exploratory, the findings of the present study would permit some specific hypotheses to be made. Most notably, one would expect that the item measuring infant's ability to disengage visual attention would *not* be a good predictor of ASD outcomes in this group. In order to better understand heterogeneity in ASD infancy markers, it would be particularly informative to consider whether and how patterns of impairment across AOSI items differ between high-risk infant groups.

More generally, future research would benefit from considering how specific biological and psychosocial vulnerability factors associated with VP birth relate to both early behavioural markers for ASD and social communication outcomes in this group. Though beyond the scope of the present project, it would be possible to explore some of these relationships within the current data set. Not only is a rich background of data regarding neonatal morbidity available for all VP infants who have taken part, but the majority of these

infants also completed EEG and MRI imaging during the neonatal period. It would be of significant interest to explore whether individual differences in brain structure (e.g. white matter volumes), general indicators of early morbidity (e.g. length of stay in specialist care) or more specific disease profiles could explain within-group variability in ASD related outcomes. Evidence is starting to emerge regarding early brain correlates associated with autism spectrum symptomatology in VP/EP groups, including links between poor brain growth during the neonatal period and later ASD diagnosis (Padilla et al., 2015). It may be that early brain markers *in combination with* behavioural markers, assessed for instance using the AOSI, will be more powerful than either in isolation as a means of predicting ASD outcomes.

Further research may also seek to explore how biological risk factors (e.g. brain abnormalities) and psychosocial (e.g. exposure to stressors; parent-infant interaction) factors interrelate and contribute to emerging difficulties with social communication. Recent research has highlighted the potentially critical interplay between psychosocial, biological and early behavioural risk factors in high-risk preterm groups, and the potential for imaging research to further this understanding (Dudova et al., 2014). For example, Smith and colleagues (2011) found that number of stressors to which an infant was exposed during their stay in the NICU was directly associated with decreased frontal and parietal brain width, and functional connectivity in the temporal lobes – as well as abnormal motor behavior during neurobehavioral examinations. Closer

examination of how environmental factors and very early brain development interrelate and contribute towards emerging ASD symptomology would therefore be an important area for further research.

Dissemination

The findings and implications of this research have been disseminated to the local research group and will also be shared in the form an oral presentation to a broader network of researchers interested in early ASD detection (the BASIS¹¹ Network). A brief and accessible summary of findings will also be shared with families who have taken part and interested members of the public via the study website. Possible avenues for submission to a peer reviewed journal will be explored.

Concluding Statements

The principal focus of this study was to investigate early markers for ASD in a VP sample. The pattern of results suggests that a structured behavioural assessment developed to measure early signs of ASD symptomatology in other high-risk infant groups (Bryson et al., 2008) may show some generalisability to preterm samples. The AOSI detected impairments in VP infants relative to controls at both 6 and 12 months, and it appeared that these impairments were independent of developmental delay. The AOSI also

¹¹ British Autism Study of Infant Siblings (<http://basisnetwork.org>)

showed some associations with ASD related outcomes at 30-34 months in this clinical group.

In contrast, two specific early infancy markers that have been linked to ASD outcomes in infants at high familial risk for ASD (attention disengagement and atypical temperament profiles) appeared *not* to be useful indicators of ASD outcomes in this VP sample. On an empirical level, these findings highlight that conclusions drawn from prospective research with infant siblings should not be assumed to be generalisable to other high-risk groups. Further research with larger longitudinal samples, longer-term follow-ups and a more fine-grained examination of individual risk factors would provide greater insight into early trajectories associated with ASD symptomology in VP groups. Further exploration of the AOSI as an early clinical screening tool in different high-risk groups would be especially beneficial.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Arlington: American Psychiatric Publishing.
- Anderson, P. J., De Luca, C. R., Hutchinson, E., Roberts, G., & Doyle, L. W. (2010). Underestimation of developmental delay by the new Bayley-III Scale. *Archives of Pediatrics & Adolescent Medicine*, *164*(4), 352–356.
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210–215.
- Baron, I., & Rey-Casserly, C. (2010). Extremely Preterm Birth Outcome: A Review of Four Decades of Cognitive Research. *Neuropsychology Review*, *20*(4), 430–452.
- Bayley, N. (2006). *Bayley scales of infant and toddler development: Bayley-III*. Harcourt Assessment, Psych. Corporation.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307.
- Bedford, R., Pickles, A., Gliga, T., Elsabbagh, M., Charman, T., & Johnson, M. H. (2014). Additive effects of social and non-social attention during infancy relate to later autism spectrum disorder. *Developmental Science*, *17*(4), 612–620.
- Behrman, R. E., & Butler, A. S. (2007). Sociodemographic and community factors contributing to preterm birth. In *Institute of Medicine (US) Committee on Understanding Premature Birth and Assuring Healthy Outcomes*. Washington, DC.: National Academies Press (US).
- Bhutta, A. T., Cleves, M. a, Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA : The Journal of the American Medical Association*, *288*(6), 728–737.
- Bloom, L., & Capatides, J. B. (1987). Expression of affect and the emergence of language. *Child Development*, 1513–1522.
- Böhm, B., Smedler, A., & Forssberg, H. (2004). Impulse control, working memory and other executive functions in preterm children when starting school. *Acta Paediatrica*, *93*(10), 1363–1371.
- Bryson, S. E., Rogers, S. J., & Fombonne, E. (2003). In review. Autism Spectrum Disorders: Early Detection, Intervention, Education, and Psychopharmacological Management. *Canadian Journal of Psychiatry*, *48*, 506–516.

- Bryson, S. E., Zwaigenbaum, L., Mcdermott, C., Rombough, V., & Brian, J. (2008). The Autism Observation Scale for Infants: scale development and reliability data. *Journal of Autism and Developmental Disorders*, *38*(4), 731–738.
- Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of autism spectrum disorders in the United Kingdom and the United States. *JAMA Pediatrics*, *168*(8), 721–728.
- Buss, A. H., & Plomin, R. (2014). *Temperament (PLE: Emotion): Early Developing Personality Traits* (Vol. 3). Psychology Press.
- Camarata, S. (2014). Early identification and early intervention in autism spectrum disorders: Accurate and effective? *International Journal of Speech-Language Pathology*, *16*(1), 1–10.
- Carey, K. B., Carey, M. P., Maisto, S. A., & Henson, J. M. (2004). Temporal stability of the timeline followback interview for alcohol and drug use with psychiatric outpatients. *Journal of Studies on Alcohol*, *65*(6), 774–781.
- Carter, M. T., & Scherer, S. W. (2013). Autism spectrum disorder in the genetics clinic: a review. *Clinical Genetics*, *83*(5), 399–407.
- Casalin, S., Luyten, P., Vliegen, N., & Meurs, P. (2012). The structure and stability of temperament from infancy to toddlerhood: A one-year prospective study. *Infant Behavior and Development*, *35*(1), 94–108.
- Chandler, S., Charman, T., Baird, G., Simonoff, E., Loucas, T., Meldrum, D., ... Pickles, A. (2007). Validation of the social communication questionnaire in a population cohort of children with autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(10), 1324–1332.
- Charman, T., & Baird, G. (2002). Practitioner Review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry*, *43*(3), 289–305.
- Chen, X., Ender, P., Mitchell, M., & Wells, C. (2003). Regression with SPSS. Retrieved April 15, 2016, from <http://www.ats.ucla.edu/stat/spss/webbooks/reg/default.htm>
- Clifford, S. M., Hudry, K., Elsabbagh, M., Charman, T., Johnson, M. H., & BASIS. (2013). Temperament in the first 2 years of life in infants at high-risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(3), 673–686.
- Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., ... Wheelwright, S. (1999). Autism spectrum disorders at 20 and 42 months of age: stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry*, *40*(5), 719–732.

- Davis, T. E., Fodstad, J. C., Jenkins, W. S., Hess, J. A., Moree, B. N., Dempsey, T., & Matson, J. L. (2010). Anxiety and avoidance in infants and toddlers with autism spectrum disorders: Evidence for differing symptom severity and presentation. *Research in Autism Spectrum Disorders, 4*(2), 305–313.
- Dawson, G., Rogers, S., Munson, J., 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.
- Day, K. L., Van Lieshout, R. J., Vaillancourt, T., & Schmidt, L. A. (2015). Peer victimization in survivors of premature birth and low birth weight: Review and recommendations. *Aggression and Violent Behavior, 25*, 259–265.
- De Groote, I., Roeyers, H., & Warreyn, P. (2006). Social-Communicative Abilities in Young High-Risk Preterm Children. *Journal of Developmental and Physical Disabilities, 18*(2), 183–200.
- De Jong, M., Verhoeven, M., & van Baar, A. L. (2012). School outcome, cognitive functioning, and behaviour problems in moderate and late preterm children and adults: a review. In *Seminars in Fetal and Neonatal Medicine* (Vol. 17, pp. 163–169). Elsevier.
- del Rosario, M., Gillespie-Lynch, K., Johnson, S., Sigman, M., & Hutman, T. (2014). Parent-reported temperament trajectories among infant siblings of children with autism. *Journal of Autism and Developmental Disorders, 44*(2), 381–393.
- Dixon Jr, W. E., & Smith, P. H. (2000). Links between early temperament and language acquisition. *Merrill-Palmer Quarterly (1982-), 46*(3), 417–440.
- Dudova, I., Kasparova, M., Markova, D., Zemankova, J., Beranova, S., Urbanek, T., & Hrdlicka, M. (2014). Screening for autism in preterm children with extremely low and very low birth weight. *Neuropsychiatric Disease and Treatment, 10*, 277.
- Durkin, M. S., Maenner, M. J., Meaney, F. J., Levy, S. E., DiGuseppi, C., Nicholas, J. S., ... Schieve, L. A. (2010). Socioeconomic inequality in the prevalence of autism spectrum disorder: evidence from a US cross-sectional study. *PLoS One, 5*(7), e11551.
- Dyet, L. E., Kennea, N., Counsell, S. J., Maalouf, E. F., Ajayi-Obe, M., Duggan, P. J., ... Herlihy, A. H. (2006). Natural history of brain lesions in extremely preterm infants studied with serial magnetic resonance imaging from birth and neurodevelopmental assessment. *Pediatrics, 118*(2), 536–548.
- Eaton, W. W., Mortensen, P. B., Thomsen, P. H., & Frydenberg, M. (2001). Obstetric Complications and Risk for Severe Psychopathology in

- Childhood. *Journal of Autism and Developmental Disorders*, 31(3), 279–285.
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.
- Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., Johnson, M. H., & BASIS. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry*, 74(3), 189–194.
- Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., Johnson, M. H., & BASIS. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research*, 251, 147–154.
- Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., ... Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, 50, 637–642.
- Escalona, S. K. (1982). Babies at double hazard: early development of infants at biologic and social risk. *Pediatrics*, 70(5), 670–676.
- Espy, K. A., Stalets, M. M., McDiarmid, M. M., Senn, T. E., Cwik, M. F., & Hamby, A. (2003). Executive Functions in Preschool Children Born Preterm: Application of Cognitive Neuroscience Paradigms. *Child Neuropsychology*, 8(2), 83–92.
- Estes, A., Munson, J., Rogers, S. J., Greenson, J., Winter, J., & Dawson, G. (2015). Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(7), 580–587.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.
- Field, M. J., & Behrman, R. E. (2004). *The necessity and challenges of clinical research involving children*. National Academies Press (US).
- Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. *Journal of Autism and Developmental Disorders*, 33(4), 365–382.
- Foster-Cohen, S., Edgin, J. O., Champion, P. R., & Woodward, L. J. (2007). Early delayed language development in very preterm infants: Evidence from the MacArthur-Bates CDI. *Journal of Child Language*, 34(03), 655–675.
- Frith, U., & Happé, F. (2005). Autism spectrum disorder. *Current Biology*, 15(19), 786–790.

- Gammer, I., Bedford, R., Elsabbagh, M., Garwood, H., Pasco, G., Tucker, L., ... Charman, T. (2015). Behavioural markers for autism in infancy: Scores on the Autism Observational Scale for Infants in a prospective study of at-risk siblings. *Infant Behavior and Development*, 38(0), 107–115.
- Garon, N., Bryson, S. E., Zwaigenbaum, L., Smith, I. M., Brian, J., Roberts, W., & Szatmari, P. (2009). Temperament and its relationship to autistic symptoms in a high-risk infant sib cohort. *Journal of Abnormal Child Psychology*, 37(1), 59–78.
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, 26(1), 64–86.
- Girgis, R. R., Minshew, N. J., Melhem, N. M., Nutche, J. J., Keshavan, M. S., & Hardan, A. Y. (2007). Volumetric alterations of the orbitofrontal cortex in autism. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 31(1), 41–45.
- Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *Lancet*, 371(9606), 75–84.
- Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., ... McCall, R. B. (1987). Roundtable: What is temperament? Four approaches. *Child Development*, 58(2), 505–529.
- Grandepeesheh, D., Dixon, D. R., Tarbox, J., Kaplan, A. M., & Wilke, A. E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 3(4), 1014–1022.
- Grunau, R. E., Haley, D. W., Whitfield, M. F., Weinberg, J., Yu, W., & Thiessen, P. (2007). Altered basal cortisol levels at 3, 6, 8 and 18 months in infants born at extremely low gestational age. *The Journal of Pediatrics*, 150(2), 151–156.
- Gurka, M. J., LoCasale-Crouch, J., & Blackman, J. A. (2010). Long-term Cognition, Achievement, Socioemotional, and Behavioral Development of Healthy Late-Preterm Infants. *Archives of Pediatrics & Adolescent Medicine*, 164(6), 525–532.
- Hack, M., Youngstrom, E. A., Cartar, L., Schluchter, M., Taylor, H. G., Flannery, D., ... Borawski, E. (2004). Behavioral Outcomes and Evidence of Psychopathology Among Very Low Birth Weight Infants at Age 20 Years. *Pediatrics*, 114(4), 932–940.
- Hanson, E., Cerban, B. M., Slater, C. M., Caccamo, L. M., Bacic, J., & Chan, E. (2013). Brief report: Prevalence of attention deficit/hyperactivity disorder among individuals with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(6), 1459–1464.

- Henderson, H. A., & Wachs, T. D. (2007). Temperament theory and the study of cognition–emotion interactions across development. *Developmental Review, 27*(3), 396–427.
- Hitzert, M. M., Van Braeckel, K. N. J. A., Bos, A. F., Hunnius, S., & Geuze, R. H. (2014). Early Visual Attention in Preterm and Fullterm Infants in Relation to Cognitive and Motor Outcomes at School Age: An Exploratory Study. *Frontiers in Pediatrics, 2*(106), 1–13.
- Hood, B. M. (1995). Shifts of visual attention in the human infant: A neuroscientific approach. *Advances in Infancy Research, 9*, 163–216.
- Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45*(2), 212–229.
- Hunnius, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy, 6*(2), 231–255.
- Hurwitz, S., & Yirmiya, N. (2014). Autism Diagnostic Observation Schedule (ADOS) and Its Uses in Research and Practice. In *Comprehensive Guide to Autism* (pp. 345–353). Springer.
- Hutchinson, E. A., De Luca, C. R., Doyle, L. W., Roberts, G., Anderson, P. J., & Group, V. I. C. S. (2013). School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics, 131*(4), 1053–1061.
- Indredavik, M. S., Vik, T., Heyerdahl, S., Kulseng, S., Fayers, P., & Brubakk, A.-M. (2004). Psychiatric symptoms and disorders in adolescents with low birth weight. *Archives of Disease in Childhood - Fetal and Neonatal Edition, 89*(5), 445–450. <http://doi.org/10.1136/adc.2003.038943>
- Isaacs, E. B., Edmonds, C. J., Chong, W. K., Lucas, A., Morley, R., & Gadian, D. G. (2004). Brain morphometry and IQ measurements in preterm children. *Brain, 127*(12), 2595–2607.
- Iwata, S., Nakamura, T., Hizume, E., Kihara, H., Takashima, S., Matsuishi, T., & Iwata, O. (2012). Qualitative brain MRI at term and cognitive outcomes at 9 years after very preterm birth. *Pediatrics, 129*(5), e1138–e1147.
- Johnson, M. H. (1990). Cortical Maturation and the Development of Visual Attention in Early Infancy. *Journal of Cognitive Neuroscience, 2*(2), 81–95.
- Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., De Haan, M., ... Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology, 17*(03), 599–619.

- Johnson, M. H., Posner, M., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Cognitive Neuroscience, Journal of*, 3(4), 335–344.
- Johnson, S. (2015). Understanding social development following very preterm birth. *Developmental Medicine and Child Neurology*, 57(10), 890–890.
- Johnson, S., Hennessy, E., Hollis, C., Kochhar, P., Wolke, D., & Marlow, N. (2009). Screening for Autism Spectrum Disorders in Extremely Preterm Children. *Acta Paediatrica*, 98, 163–163.
- Johnson, S., Hollis, C., Hennessy, E., Kochhar, P., Wolke, D., & Marlow, N. (2011). Screening for autism in preterm children: diagnostic utility of the Social Communication Questionnaire. *Archives of Disease in Childhood*, 96(1), 73–77.
- Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., & Marlow, N. (2010a). Autism spectrum disorders in extremely preterm children. *The Journal of Pediatrics*, 156(4), 525–531.
- Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., & Marlow, N. (2010b). Psychiatric Disorders in Extremely Preterm Children: Longitudinal Finding at Age 11 Years in the EPICure Study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(5), 453–463.
- Johnson, S., & Marlow, N. (2011). Preterm Birth and Childhood Psychiatric Disorders. *Pediatric Research*, 69(5), 11R–18R.
- Jones, E. J. H., 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.
- Kaarensen, P. I., Rønning, J. A., Ulvund, S. E., & Dahl, L. B. (2006). A Randomized, Controlled Trial of the Effectiveness of an Early-Intervention Program in Reducing Parenting Stress After Preterm Birth. *Pediatrics*, 118(1), e9–e19.
- Kagan, J., Snidman, N., Arcus, D., & Reznick, J. S. (1994). *Galen's prophecy: Temperament in human nature*. Basic Books.
- Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, 2(10), 389–398.
- Kozlowski, A. M., Matson, J. L., Belva, B., & Rieske, R. (2012). Feeding and sleep difficulties in toddlers with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 6(1), 385–390.
- Kuban, K. C. K., Joseph, R. M., O'Shea, T. M., Allred, E. N., Heeren, T., Douglass, L., ... Hirtz, D. (2016). Girls and Boys Born before 28 Weeks

- Gestation: Risks of Cognitive, Behavioral, and Neurologic Outcomes at Age 10 Years. *The Journal of Pediatrics*, 173, 69–75.
- Kuban, K. C. K., O'Shea, T. M., Allred, E. N., Tager-Flusberg, H., Goldstein, D. J., & Leviton, A. (2009). Positive Screening on the Modified Checklist for Autism in Toddlers (M-CHAT) in Extremely Low Gestational Age Newborns. *The Journal of Pediatrics*, 154(4), 535–540.e1.
- Kuzniewicz, M. W., Wi, S., Qian, Y., Walsh, E. M., Armstrong, M. A., & Croen, L. A. (2014). Prevalence and neonatal factors associated with autism spectrum disorders in preterm infants. *The Journal of Pediatrics*, 164(1), 20–25.
- Landry, R., & Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry*, 45(6), 1115–1122.
- Langerock, N., de Jonge, L. van H., Graz, M. B., Hüppi, P. S., Tolsa, C. B., & Barisnikov, K. (2013). Emotional reactivity at 12 months in very preterm infants born at < 29 weeks of gestation. *Infant Behavior and Development*, 36(3), 289–297.
- Larsson, H. J., Eaton, W. W., Madsen, K. M., Vestergaard, M., Olesen, A. V., Agerbo, E., ... Mortensen, P. B. (2005). Risk Factors for Autism: Perinatal Factors, Parental Psychiatric History, and Socioeconomic Status. *American Journal of Epidemiology*, 161(10), 916–925.
- Lauritsen, M., & Ewald, H. (2001). The genetics of autism. *Acta Psychiatrica Scandinavica*, 103(6), 411–427.
- Limperopoulos, C., Bassan, H., Gauvreau, K., Robertson, R. L., Sullivan, N. R., Benson, C. B., ... du Plessis, A. J. (2007). Does cerebellar injury in premature infants contribute to the high prevalence of long-term cognitive, learning, and behavioral disability in survivors? *Pediatrics*, 120(3), 584–593.
- Limperopoulos, C., Bassan, H., Sullivan, N. R., Soul, J. S., Robertson, R. L., Moore, M., ... du Plessis, A. J. (2008). Positive screening for autism in ex-preterm infants: Prevalence and risk factors. *Pediatrics*, 121(4), 758–765.
- Limperopoulos, C., Robertson, R. L., Sullivan, N. R., Bassan, H., & du Plessis, A. J. (2009). Cerebellar Injury in Term Infants: Clinical Characteristics, Magnetic Resonance Imaging Findings, and Outcome. *Pediatric Neurology*, 41(1), 1–8.
- Little, T. D. (2013). *The Oxford handbook of quantitative methods*. New York: Oxford University Press.
- Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A.

- (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, 63(6), 694–701.
- Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., ... Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule: ADOS-2*. Western Psychological Services Los Angeles, CA.
- Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal*, 14(3), 234–252.
- Luyster, R., Guthrie, W., Gotham, K., Risi, S., DiLavore, P., & Lord, C. (2009). The Autism Diagnostic Observation Schedule—Toddler module: Preliminary findings using a modified version of the ADOS. In *International meeting for autism research* (pp. 15–17).
- Luyster, R. J., Kuban, K. C. K., O’Shea, T. M., Paneth, N., Allred, E. N., Leviton, A., & For, E. S. investigators. (2011). The Modified Checklist for Autism in Toddlers in extremely low gestational age newborns: individual items associated with motor, cognitive, vision and hearing limitations. *Paediatric and Perinatal Epidemiology*, 25(4), 366–376.
- MacKay, D. F., Smith, G. C. S., Dobbie, R., & Pell, J. P. (2010). Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren. *PLoS Medicine*, 7(6), e1000289.
- Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders*, 42(7), 1304–1313.
- Mannion, A., & Leader, G. (2013). Comorbidity in autism spectrum disorder: A literature review. *Research in Autism Spectrum Disorders*, 7(12), 1595–1616.
- Marlow, N., Wolke, D., Bracewell, M. A., & Samara, M. (2005). Neurologic and developmental disability at six years of age after extremely preterm birth. *New England Journal of Medicine*, 352(1), 9–19.
- Marret, S., Marchand-Martin, L., Picaud, J.-C., Hascoët, J.-M., Arnaud, C., Rozé, J.-C., ... Ancel, P.-Y. (2013). Brain injury in very preterm children and neurosensory and cognitive disabilities during childhood: the EIPAGE cohort study, 8(5), e62683.

- Martin, R. P., Wisenbaker, J., Baker, J., & Huttunen, M. O. (1997). Gender differences in temperament at six months and five years. *Infant Behavior and Development*, 20(3), 339–347.
- Maxwell, S. E. (2000). Sample size and multiple regression analysis. *Psychological Methods*, 5(4), 434.
- McConnell, B. A., & Bryson, S. E. (2005). Visual attention and temperament: Developmental data from the first 6 months of life. *Infant Behavior and Development*, 28(4), 537–544.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 40–48.
- Miller, S. P., Ferriero, D. M., Leonard, C., Piecuch, R., Glidden, D. V., Partridge, J. C., ... Barkovich, A. J. (2005). Early brain injury in premature newborns detected with magnetic resonance imaging is associated with adverse early neurodevelopmental outcome. *The Journal of Pediatrics*, 147(5), 609–616.
- Moore, T., Johnson, S., Hennessy, E., & Marlow, N. (2012). Screening for autism in extremely preterm infants: problems in interpretation. *Developmental Medicine & Child Neurology*, 54(6), 514–520.
- Moore, V., & Goodson, S. (2003). How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. *Autism*, 7(1), 47–63.
- Nadeau, L., Tessier, R., Boivin, M., Lefebvre, F., & Robaey, P. (2003). Extremely Premature and Very Low Birthweight Infants: A Double Hazard Population? *Social Development*, 12(2), 235–248.
- NICE. (2011). Autism in under 19s: recognition, referral and diagnosis. Retrieved May 13, 2016, from <https://www.nice.org.uk/guidance/cg128>
- NICE. (2015). NICE Guideline scope: Developmental follow-up of preterm babies final scope (Aug 2015). Retrieved April 23, 2016, from <https://www.nice.org.uk/guidance/gid-cgwave0752/resources/developmental-followup-of-preterm-babies-final-scope2>
- Noel, M., Peterson, C., & Jesso, B. (2008). The relationship of parenting stress and child temperament to language development among economically disadvantaged preschoolers. *Journal of Child Language*, 35(04), 823–843.
- O'Shea, T. M., Allred, E. N., Dammann, O., Hirtz, D., Kuban, K. C. K., Paneth, N., ... Investigators, E. S. (2009). The ELGAN study of the brain and related disorders in extremely low gestational age newborns. *Early*

Human Development, 85(11), 719–725.

- Oakes, M. (2006). Measuring Socioeconomic Status. Retrieved May 27, 2016, from http://www.esourceresearch.org/Portals/0/Uploads/Documents/Public/Oakes_FullChapter.pdf
- Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*, 24(3), 247–257.
- Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research*, 47(17), 2272–2280.
- Padilla, N., Eklöf, E., Mårtensson, G. E., Bölte, S., Lagercrantz, H., & Ådén, U. (2015). Poor Brain Growth in Extremely Preterm Neonates Long Before the Onset of Autism Spectrum Disorder Symptoms. *Cerebral Cortex, E-Publicat*, bhv300.
- Pallant, J. (2007). SPSS survival manual: A step-by-step guide to data analysis using SPSS version 15. *Nova lorque: McGraw Hill*.
- Parker, S., Greer, S., & Zuckerman, B. (1988). Double jeopardy: the impact of poverty on early child development. *Pediatric Clinics of North America*, 35(6), 1227–1240.
- Perlman, J. M. (2001). Neurobehavioral deficits in premature graduates of intensive care - Potential medical and neonatal environmental risk factors. *Pediatrics*, 108(6), 1339–1348.
- Perlman, J. M. (2003). The genesis of cognitive and behavioral deficits in premature graduates of intensive care. *Minerva Pediatrica*, 55(2), 89–101.
- Perry, A., Flanagan, H. E., Geier, J. D., & Freeman, N. L. (2009). Brief report: The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism and Developmental Disorders*, 39(7), 1066–1078.
- Picciolini, O., Squarza, C., Fontana, C., Giannì, M. L., Cortinovis, I., Gangi, S., ... Mosca, F. (2015). Neurodevelopmental outcome of extremely low birth weight infants at 24 months corrected age: a comparison between Griffiths and Bayley Scales. *BMC Pediatrics*, 15(1), 1–9.
- Pilling, S., Baron-Cohen, S., Megnin-Viggars, O., Lee, R., Taylor, C., & Group, G. D. (2012). Recognition, referral, diagnosis, and management of adults with autism: summary of NICE guidance. *BMJ*, 344, e4082.
- Pineda, R. G., Neil, J., Dierker, D., Smyser, C. D., Wallendorf, M., Kidokoro,

- H., ... Mathur, A. M. (2014). Alterations in brain structure and neurodevelopmental outcome in preterm infants hospitalized in different neonatal intensive care unit environments. *The Journal of Pediatrics*, *164*(1), 52–60. e2.
- Pinto-Martin, J., & Levy, S. (2004). Early diagnosis of autism spectrum disorders. *Current Treatment Options in Neurology*, *6*(5), 391–400.
- Posner, M. I. (2014). Guides to the study of attention. In K. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention* (p. 1). OUP Oxford.
- Posner, M. I., Walker, J. A., Friedrich, F. A., & Rafal, R. D. (1987). How do the parietal lobes direct covert attention? *Neuropsychologia*, *25*(1), 135–145.
- Pottie, C. G., & Ingram, K. M. (2008). Daily stress, coping, and well-being in parents of children with autism: a multilevel modeling approach. *Journal of Family Psychology*, *22*(6), 855–864.
- Pritchard, M. A., de Dassel, T., Beller, E., Bogossian, F., Johnston, L., Paynter, J., ... Scott, J. (2016). Autism in Toddlers Born Very Preterm. *Pediatrics*, *137*(2), 1–8.
- Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. *Advances in Research on Temperament*, 165–182.
- Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior and Development*, *29*(3), 386–401.
- Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and Assessment of Short and Very Short Forms of the Infant Behavior Questionnaire–Revised. *Journal of Personality Assessment*, *96*(4), 445–458.
- Rai, D., Lewis, G., Lundberg, M., Araya, R., Svensson, A., Dalman, C., ... Magnusson, C. (2012). Parental socioeconomic status and risk of offspring autism spectrum disorders in a Swedish population-based study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(5), 467–476.
- Rajulton, F. (2001). The fundamentals of longitudinal research: An overview. *Canadian Studies in Population*, *28*(2), 169–185.
- Rees, G. (2009). Visual attention: the thalamus at the centre? *Current Biology*, *19*(5), R213–R214.
- Ritchie, K., Bora, S., & Woodward, L. J. (2015). Social development of children born very preterm: a systematic review. *Developmental Medicine*

and *Child Neurology*, 57(10), 899–918.

- Robertson, C. M. T., Hendson, L., Biggs, W. S. G., & Acton, B. V. (2010). Application of the Flynn effect for the Bayley III Scales. *Archives of Pediatrics & Adolescent Medicine*, 164(11), 1072–1073.
- Robins, D., Fein, D., Barton, M., & Green, J. (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., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., ... Dawson, G. (2012). Effects of a brief Early Start Denver Model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(10), 1052–1065.
- Roncadin, C., Jilderda, S., Brian, J., Roberts, W., Smit, M., Bryson, S., ... Zwaigenbaum, L. (2011). Behavioral Signs of Autism In Premature Infants: Findings From the Autism Observation Scale for Infants and Autism Diagnostic Observation Schedule. In *International meeting for autism research*. San Diego, USA.
- Roncadin, C., Rourke, M., Jilderda, S., Brian, J., Bryson, S. E., Niccols, A., ... Zwaigenbaum, L. (2010). Behavioral Profiles of Premature Infants: Similarities and Differences to Infant Siblings Later Diagnosed with Autism. In *International meeting for autism research*. Philadelphia, USA.
- Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. *Temperament: Individual Differences at the Interface of Biology and Behavior*, 83–116.
- Rothbart, M. K., Ellis, L. K., & Posner, M. I. (2004). Temperament and self-regulation. *Handbook of Self-Regulation: Research, Theory, and Applications*, 2, 284–299.
- Rothbart, M. K., & Mauro, J. A. (1990). Questionnaire approaches to the study of infant temperament. *Individual Differences in Infancy: Reliability, Stability and Prediction*, 411–429.
- Russell, J. E. (1997). *Autism as an executive disorder*. Oxford University Press.
- Rutter, M., & Thapar, A. (2014). Genetics of autism spectrum disorder. In F. R. Volkmar, R. Paul, K. Pelphrey, & S. Roger (Eds.), *Handbook of autism* (4th ed.). Hoboken, NJ: John Wiley.
- Sacrey, L.-A. R., Armstrong, V. L., Bryson, S. E., & Zwaigenbaum, L. (2014). Impairments to visual disengagement in autism spectrum disorder: A review of experimental studies from infancy to adulthood. *Neuroscience &*

Biobehavioral Reviews, 47, 559–577.

- Saigal, S., & Doyle, L. W. (2008). An overview of mortality and sequelae of preterm birth from infancy to adulthood. *The Lancet*, 371(9608), 261–269.
- Salley, B. J., & Dixon Jr, W. E. (2007). Temperamental and joint attentional predictors of language development. *Merrill-Palmer Quarterly (Wayne State University. Press)*, 53(1), 131.
- Sameroff, A. J., Chandler, M. J., & Horowitz, F. D. (1975). Review of child development research. *Review of Child Development Research*, 4.
- Sayal, K., Heron, J., Maughan, B., Rowe, R., & Ramchandani, P. (2014). Infant temperament and childhood psychiatric disorder: longitudinal study. *Child: Care, Health and Development*, 40(2), 292–297.
- Schendel, D., & Bhasin, T. K. (2008). Birth Weight and Gestational Age Characteristics of Children With Autism, Including a Comparison With Other Developmental Disabilities. *Pediatrics*, 121(6), 1155–1164. <http://doi.org/10.1542/peds.2007-1049>
- Schmidt, B., Asztalos, E., & Roberts, R. (2003). Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months: Results from the trial of indomethacin prophylaxis in preterms. *JAMA*, 289(9), 1124–1129.
- Seifer, R., Sameroff, A. J., Barrett, L. C., & Krafchuk, E. (1994). Infant temperament measured by multiple observations and mother report. *Child Development*, 65(5), 1478–1490.
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(8), 921–929.
- Skuse, D. (2004). The developmental, dimensional and diagnostic interview (3di): A novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 43(5), 548–558.
- Smith, G. C., Gutovich, J., Smyser, C., Pineda, R., Newnham, C., Tjoeng, T. H., ... Inder, T. (2011). Neonatal intensive care unit stress is associated with brain development in preterm infants. *Annals of Neurology*, 70(4), 541–549.
- Smith, I. C., Reichow, B., & Volkmar, F. R. (2015). The Effects of DSM-5 Criteria on Number of Individuals Diagnosed with Autism Spectrum Disorder: A Systematic Review. *Journal of Autism and Developmental Disorders*, 45(8), 2541–2552.

- Smith, L. K., Draper, E. S., Manktelow, B. N., Dorling, J. S., & Field, D. J. (2007). Socioeconomic inequalities in very preterm birth rates. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 92(1), F11–F14.
- Soper, D. S. (2016). Post-hoc Statistical Power Calculator for Multiple Regression [Software]. Retrieved May 2, 2016, from <http://www.danielsoper.com/statcalc/calculator.aspx?id=1>
- Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., & Weiss, B. H. (1999). Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry*, 40(02), 219–226.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Thomson/Brooks/Cole.
- Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2007). *Using Multivariate Statistics (Fifth Edition)*. Pearson.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. Brunner/Mazel.
- Treyvaud, K., Ure, A., Doyle, L. W., Lee, K. J., Rogers, C. E., Kidokoro, H., ... Anderson, P. J. (2013). Psychiatric outcomes at age seven for very preterm children: rates and predictors. *Journal of Child Psychology and Psychiatry*, 54(7), 772–779.
- van de Weijer-Bergsma, E., Wijnroks, L., & Jongmans, M. J. (2008). Attention development in infants and preschool children born preterm: A review. *Infant Behavior and Development*, 31(3), 333–351.
- VanVoorhis, C. R. W., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology*, 3(2), 43–50.
- Vicari, S., Caravale, B., Carlesimo, G. A., Casadei, A. M., & Allemand, F. (2004). Spatial working memory deficits in children at ages 3-4 who were low birth weight, preterm infants. *Neuropsychology*, 18(4), 673–678.
- Volkmar, F. R., & McPartland, J. C. (2014). From Kanner to DSM-5: autism as an evolving diagnostic concept. *Annual Review of Clinical Psychology*, 10, 193–212.
- Waddington, H., van der Meer, L., & Sigafos, J. (2016). Effectiveness of the Early Start Denver Model: a Systematic Review. *Review Journal of Autism and Developmental Disorders*, 1–14.
- Wainwright, A., & Bryson, S. E. (2005). The development of endogenous orienting: Control over the scope of attention and lateral asymmetries. *Developmental Neuropsychology*, 27(2), 237–255.

- Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-VanderWeele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*, *127*(5), e1303–e1311.
- Wass, S., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology*, *21*(18), 1543–1547.
- Wilson, S. L., & Craddock, M. M. (2004). Review: Accounting for prematurity in developmental assessment and the use of age-adjusted scores. *Journal of Pediatric Psychology*, *29*(8), 641–649.
- Woodward, L. J., Anderson, P. J., Austin, N. C., Howard, K., & Inder, T. E. (2006). Neonatal MRI to predict neurodevelopmental outcomes in preterm infants. *New England Journal of Medicine*, *355*(7), 685–694.
- Yaari, M., Yitzhak, N., Harel, A., Friedlander, E., Bar-Oz, B., Eventov-Friedman, S., ... Yirmiya, N. (2016). Stability of early risk assessment for autism spectrum disorder in preterm infants. *Autism, E-Publicat.*
- Yirmiya, N., & Charman, T. (2010). The prodrome of autism: early behavioral and biological signs, regression, peri- and post-natal development and genetics. *Journal of Child Psychology and Psychiatry*, *51*(4), 432–458.
- Zachor, D. A., Ben-Itzhak, E., Rabinovich, A.-L., & Lahat, E. (2007). Change in autism core symptoms with intervention. *Research in Autism Spectrum Disorders*, *1*(4), 304–317.
- Zeitlin, J., Ancel, P.-Y., Larroque, B., Kaminski, M., & EPIPAGE. (2004). Fetal sex and indicated very preterm birth: results of the EPIPAGE study. *American Journal of Obstetrics and Gynecology*, *190*(5), 1322–1325.
- Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of autism spectrum disorders. *Behavioural Brain Research*, *251*, 133–146.
- 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*, 143–152.
- Zwaigenbaum, L., Thurm, A., Stone, W., Baranek, G., Bryson, S., Iverson, J., ... Sigman, M. (2007). Studying the Emergence of Autism Spectrum Disorders in High-risk Infants: Methodological and Practical Issues. *Journal of Autism and Developmental Disorders*, *37*(3), 466–480.

Appendices

Appendix 1 – NHS Ethics Approval Letter



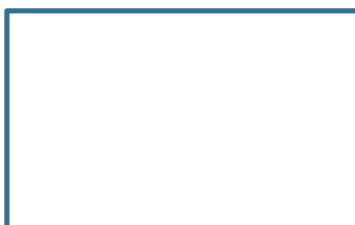
National Research Ethics Service

North West London REC 2

Royal Free Hospital NHS Trust
Royal Free Hospital
South House, Block A
Pond Street
London
NW3 2QG

Telephone: 020 7317 7718
Facsimile: 020 7794 0714

15 February 2011



Study Title:



REC reference number: very preterm birth
10/H0720/80

Thank you for your letter of 31st January, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

The Committee has not yet been notified of the outcome of any site-specific assessment (SSA) for the non-NHS research site(s) taking part in this study. The favourable opinion does not therefore apply to any non-NHS site at present. I will write to you again as soon as one Research Ethics Committee has notified the outcome of a SSA. In the meantime no study procedures should be initiated at non-NHS sites.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where the only involvement of the NHS organisation is as a Participant Identification Centre (PIC), management permission for research is not required but the R&D office should be notified of the study and agree to the organisation's involvement. Guidance on procedures for PICs is available in IRAS. Further advice should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	Version	Date
Protocol	1.0	03 September 2010
Response to Request for Further Information		
GP/Consultant Information Sheets	1.0	03 September 2010
REC application		09 September 2010
Questionnaire: Non-Validated: SEC Questionnaire	1.0	09 September 2010
Poster for Neonatal Unit	1.0	03 September 2010
Participant Information Sheet: Parent Information Leaflet	2.0	06 December 2010
Participant Information Sheet: Children Born at Term	1.0	06 December 2010
Investigator CV		06 September 2010
Participant Consent Form	1.0	03 September 2010
Evidence of insurance or indemnity		09 September 2010
Referees or other scientific critique report		
Summary/Synopsis	1	09 September 2010
Webpage for IfWH Website	1.0	03 September 2010
Letter from Funder		14 April 2010

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National

Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "*After ethical review – guidance for researchers*" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

10/H0720/80	Please quote this number on all correspondence
-------------	--

With the Committee's best wishes for the success of this project

Yours sincerely



Dr Michael Pegg
Chair

Email: Thomas.mcquillan@royalfree.nhs.uk

Enclosures: "After ethical review – guidance for researchers" [SL-AR1 for CTIMPs, SL- AR2 for other studies]

Appendix 2 – NHS Ethics Substantial Amendment Approval Letter (A)



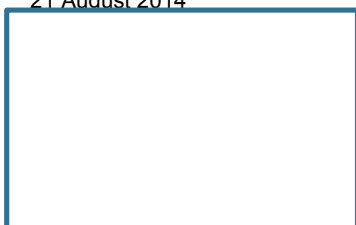
National Research Ethics Service

NRES Committee London - Hampstead

Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Tel: 0161 625 7815
Fax: 0161 625 7299

21 August 2014



Dear Professor Marlow

Study title:



g up after

REC reference:

very preterm birth

Amendment number:

10/H0720/80

Amendment date:

Amendment 3

IRAS project ID:

28 July 2014

57812

- The amendment proposes to extend recruitment until December 2016 and extend the study duration until December 2018.
- It is also proposed to include an additional assessment point when the child is two years old.
- The amendment also proposes to include a battery of MRI sequences and neuromonitoring for the participants.

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

There were no ethical issues raised.

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Notice of Substantial Amendment (non-CTIMP)	Amendment 3	28 July 2014
Other [REC Approval Letter]	13/LO/0225	12 April 2013
Other [NHS R&D Approval Letter]	14/LO/0225	10 May 2013
Participant consent form	1.0	28 July 2014
Participant information sheet [Babylab Follow Up]	1	18 July 2014
Participant information sheet	4.1	15 March 2013
Research protocol or project proposal	1.3	18 July 2014

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at <http://www.hra.nhs.uk/hra-training/>

10/H0720/80: Please quote this number on all correspondence

Yours sincerely



Signed on behalf of:
Miss Stephanie Ellis
Chair

E-mail: nrescommittee.london-hampstead@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

Copy to:

Appendix 3 – NHS Ethics Substantial Amendment Approval Letter (B)



Health Research Authority

Research Ethics Service

London - Hampstead Research Ethics Committee

Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Tel: 0207 104 8002

23 December 2015

Study title:

ing up after

REC reference: very preterm birth
10/H0720/80
Amendment number: Substantial Amendment 4
Amendment date: 27 July 2015
IRAS project ID: 57812

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

Approval was sought for the request to recruit from local nurseries. The Participant Information Sheet and Consent Form were also updated and submitted for review.

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Notice of Substantial Amendment (non-CTIMP)		27 July 2015
Other [Letter to Nursery]	1	25 July 2015
Participant consent form [Parent]	1.0	25 July 2015
Participant information sheet (PIS) [Parent]	1.0	25 July 2015

A Research Ethics Committee established by the Health Research Authority

Research protocol or project proposal	1.4	25 July 2015
---------------------------------------	-----	--------------

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at <http://www.hra.nhs.uk/hra-training/>

10/H0720/80:	Please quote this number on all correspondence
---------------------	---

Yours sincerely



**On behalf of
Miss Stephanie Ellis
Chair**

E-mail: nrescommittee.london-hampstead@nhs.net

Enclosures: List of names and professions of members who took part in the review

Copy to:


--

Appendix 4 – Royal Holloway Ethics Approval (Email)

outlook.office.com

Reply all | Delete | Junk | ...

Ref: 2015/051 **Ethics** Form Approved

 psychology.it.support@rhul.ac.uk
05/06/2015
pava067@rhul.ac.uk; Owen, Tamsin; PSY-EthicsAdmin@rhul.ac.uk; Zagefka, Hanna; Lock, Annette; uqjt005@rhul.ac.uk

Inbox

Application Details: View the form click [here](#) Revise the form click [here](#)

Applicant Name: **Charlotte Sanderson**

Application title: **Early Detection of Autism Spectrum Symptomology in Very Preterm Infants**

Comments: Approved.

0 of 0

Appendix 5 – Table of A-priori Power Calculations

Paper	Sample Characteristics	Study Focus	Test type	Group Mean (SD)		Correlation Coefficient	Sample size per group required for desired power (0.8) and alpha level (0.05)
				Group 1	Group 2		
<i>De Groot et al (2006)</i>	Preterm (<37 weeks) vs. Full term controls	Comparison of scores on the ADOS at 24 months	Between-Subjects Test of Difference (2 groups) Communication: Social Interaction:	1.16 (1.43) 1.63 (1.57)	2.88 (1.56) 3.8 (3.98)		11 16
<i>Roncadin et al (2011)</i>	Preterm (born at <37 weeks; excluded infants who later receive an ASD diagnosis)	Relationship between AOSI scores (Autism Observation Schedule for Infants) at 12 months and ADOS scores at 24 months	Correlation			0.46	35
<i>Zwaigenbaum et al (2005)</i>	Genetically high risk infants (infant siblings of children with ASD) who do and do not receive a diagnosis at 24 months	Comparison of IBQ scores at 12 months	Between-Subjects Test of Difference (2 groups)	Not reported	Not reported		Could not calculate as means were not reported. But authors report that infants later diagnosed with ASD were reported to have more "frequent and intense distress reactions to a variety of stimuli ($F_{3,56} = 4.29$; $P = 0.009$)" at 12 months. This indicates that a study sample of 30 per group would be sufficient to detect an effect at $\alpha = .05$.

Appendix 6 – Letter to Nurseries (30-34 month recruitment)

Dear Sir or Madam,

We are a team of researchers from [redacted] and we are currently carrying out a study into child development after very preterm birth. This is important research as we know that some babies born very early go on to have developmental difficulties. Our study (the [redacted]) aims to identify the earliest signs of problems, and it is hoped that this will help us to develop earlier targeted interventions for these vulnerable infants. We are writing to you today to request your help in finding willing families to take part in this research.

The [redacted] is a study that has been running for several years now. We have been recruiting babies born very preterm (i.e. at less than 32 weeks gestation) and another group of infants born normally at full term (around the time of their due dates), and tracking different aspects of their development over the first two years of life. Our group of preterm infants is now reaching 2 years of age, and we are looking for some additional healthy, term-born 2-year olds to complete some tasks as a developmental comparison.

We are writing to ask if you would be willing to pass on some information about our research to parents of children attending your nursery. If you are keen to help, we will simply ask you to share our information leaflets. The leaflet (example enclosed) contains information about the study and advises parents how to contact us if they are interested in taking part. If appropriate, we also sometimes arrange to come into nursery and speak with parents briefly to introduce ourselves and the study. The research all takes place at [redacted], we then organise everything else directly with them. We are only asking for your help in sharing information about the study.

We would be very happy to meet with you or talk further over the phone to discuss the study and how you can help further. If you are happy for us to do so, we will follow up this letter with a phone call. **If you would prefer us not to contact you, we'd be very grateful if you could complete and return the enclosed reply slip (prepaid envelope provided).**

If you would like any further information, or if you have any questions, please contact Kayleigh Day using the contact details above.

Many thanks in advance for your help.

Yours sincerely,

Name of nursery:

Are you happy for us to contact you? Yes No

If you do not reply, we will contact you to check that you have received the study information and to see if you would be interested in helping us with this research.

If you are happy to take part, you do not need to send the reply and we will contact you in the near future to talk to you about your participation.

If however, there is a specific person we would need to talk to at the nursery, we would be extremely grateful if you could send us a name and the preferred method of contact in the space below:

Name

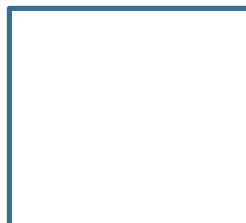
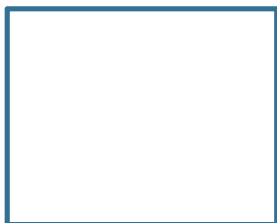
Number

Email

Please circle the preferred method of contact.

Or alternatively, please feel free to contact Kayleigh Day on either of the numbers above or by email.

Appendix 7 – 30-34 Month Only Parent Information Leaflet



THE UCH PRETERM DEVELOPMENT PROJECT *Testing children's' attention and memory at ages 2-3* Parent information leaflet

We would like to invite you to be part of our research study. To help you decide whether this is something that you would be interested in, we have put together some information to explain why this research is important and what it would involve for you and your child. If you require any further information, please contact [redacted] whose details are at the end of this leaflet.

The first part of this leaflet tells you about the study and the second part answers some frequently asked questions about research studies carried out here at [redacted]



Version 1.0 June 2015

[redacted] Preterm Development Project

For many years [redacted] we have had a very active research programme that helps make our care the best it can be. In particular, we have been studying the development of babies who are born well before their due dates, called preterm babies. As part of our normal service, we follow up all preterm babies after they leave intensive care for at least 2 years after going home, to ensure that their development is progressing nicely, and to arrange help if there are any problems.



To work out how well the measures we take are tracking different aspects of child development, we need to compare the development of babies born preterm with that of children who are born normally at full term (around the time of their due dates, after 37-41 weeks of gestation). This is why we are approaching you now, to ask if you would take part in this study as part of the full term group.

Why this is important?

This study aims to find out which aspects of very premature babies' brain growth and early development over the first few years are different from those of babies born at full term.

We know that babies who are born before 32 weeks of gestation (more than 8 weeks early) can develop perfectly normally, but for some babies, development is slow and they may need extra help to progress as well as possible. A number of studies have been conducted in older children, which have given us a better understanding of the developmental problems some of these children face later in life.

Currently, we are working hard to try to develop programmes to improve the developmental problems faced by preterm children in later childhood. However, we could be more targeted in our approach if we could identify the earliest signs of problems. If, from our research, we can detect an area that is likely to be affected in childhood, we could develop an intervention scheme specifically targeting this from the time the baby is in the neonatal unit, or soon after. By doing so, we hope to give all premature babies the very best start.

_____ have been following a group of preterm children from birth and now the cohort is reaching 2 years of age, we are looking for a group of healthy term born 2 year olds to act as a developmental comparison for the next stage of our study.

In the next part of this leaflet, we explain exactly what this would involve for you and your child and if you have any further questions, please do not hesitate to get in touch.

What does the study involve?

If you are happy to take part in the study, we would like you to come and join us in _____. During your visit, we will play a series of games and tasks that will test your child's memory, attention, language and general understandings. These tasks only take around an hour and a half to complete and we generally like to do this in a morning so everyone is bright eyed. We then tend to like to take a small break for lunch, and return in the afternoon to carry out a standardised developmental assessment from which we can provide you feedback on your child's developmental stage.

All of our young participants seem to thoroughly enjoy all of our activities and will be provided with a toy at the end of the day as a token of our appreciation.

Below, are some more details on the games and tasks we run during the day.

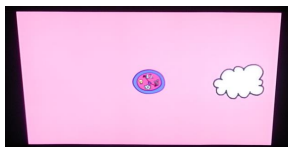
How does your child switch their attention?



We are interested in how easily your child switches his/her attention between objects and events that s/he finds interesting. The way in which children and infants visually explore their environment is very important for learning and social development.

To measure this, we will sit your child in front of a television screen that will show a series of cartoon

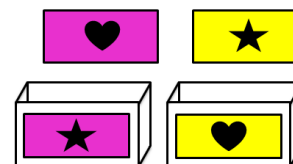
pictures. Moving objects are then randomly presented on the left and right of the cartoon. We will then measure how long it takes your child to move their attention between the animations. This task enables us to assess how flexibly children can switch their attention.



Memory and flexible understanding:

In the morning of your visit, there will be a handful of short games that we will play at a table that will assess your child's memory and more general understandings.

Between the ages of two and three, children are extremely perceptive of the world around them, but at this age they struggle to understand another person's point of



view, or be able to match a picture in two different ways. For example, if you look at the picture of the two boxes and 2 cards on the left, a child of two years may be able to sort the cards according to the colours but not according to their shapes. This type of sorting task will be one of the games we play in the morning.

The memory tasks that we run involve us hiding either stickers or a small snack (usually raisins), into various pots and in different situations and asking the child to remember where they are hidden. At the end of these tasks, your child will be able to choose a 'Well done' sticker for all their hard work.

The last task we run at the table involves testing your child's self-control. For this we ask you to bring in their favourite treat (or if you are happy, we have some small chocolate treats we can use). For this task, we place the treat under a cup and tell the child that they can eat it but only when we say, at which point we have a set delay and then they are allowed to have it. Our participants quite like this task!

How does your child process sounds?

For this part of the visit, we would like to perform an EEG recording. It is perfectly safe, and involves your child wearing a stretchy hat with small, soft sensors – like the one worn by the baby in the picture on the right, in order to see how his/her brain is functioning when listening to some sounds. This lasts no more than 10 minutes and during this time the child can watch their favourite show on a screen in front of them.



How does your child interact with you and with other adults?

In order for us to get a better understanding of your child's general social development, we would like to observe the way they interact with yourself and other adults.

During the visit, we will ask you to play with your child for a few minutes, in the same manner you normally would at home. This will show us how your child behaves when interacting with someone they know well.

We will also run a social skill assessment, where one of our research team will play with a selection of toys with your child, allowing us to see their use of language and communication abilities in a more natural setting.

We will video all our tasks on the day so that we have record of the assessment for a later date. Please see the end of the leaflet for our policy on data protection.

The standardised developmental assessment

In the afternoon, as mentioned at the beginning of this leaflet, we would like to run a standardised assessment known as the Bayley Scales of Infant and Toddler Development (Bayley-III).

This is an internationally recognised developmental assessment designed to assess all key areas of a child's development, from their language through to their motor skills. Children born preterm are assessed using this scale in the clinical follow ups, but this is not route for children who are developing without any concerns. We would therefore like to run this with our term children at this age so to have a comparison score for our children that are born preterm. As this is a clinical assessment, and has been standardised over a large group of children, we are able to provide you with a feedback report on what stage your child is at for all the different developmental areas.



This is the last task in the study and all together it should take more than 5 hours including a break for lunch. We would also be very happy to contribute to your lunch and we can cover travel expenses upon presentation of a receipt.

During your visit, we will discuss the possibility of later follow-ups with you. We would like carry out some further assessments as our cohort grows up so we can find out how our early measures are related to the child's development later in childhood. We ask your permission to do this in the consent form but initially you are only signing up to this one visit with the study.

What do we need to do now?

If you are happy to take part in our study and feel comfortable that you understand what it is we are aiming to achieve with the research, we will ask you to sign a consent form. We will do this at the beginning of your visit with us in the baby lab, so all we need from you initially are your contact details and the due date and birth date of your child in order for us to confirm that your child is eligible for the study. We are ideally looking for your child to be around 30 months of age (two and a half years) but if you are interested and your child is either side of this age range we would still love to hear from you.

Some Frequently Asked Questions – FAQ:

Are there disadvantages or risks to taking part in the study?

There are no risks associated to any of the task we run in these assessments. The EEG is perfectly safe and for each task we will only continue as long as your child is happy remains interested in it.

Are there any benefits to taking part?

This study will not be of any direct benefit to your child, but if it seems to us that your child needs extra support Prof Marlow will discuss with you how this can be provided.

How can you find out more about the study and about very preterm babies?

If you have any concerns or questions about the study please feel free to contact a member of the team who will be more than happy to provide any extra information you require. The support organisation, BLISS, has a very informative website explaining more about preterm babies. All contact information and relevant web addresses can be found at the end of this leaflet.

Where will the study take place?

The study will take place in happy to reimburse your travel costs to and from the hospital on presentation of a receipt and we will send your child a certificate to acknowledge their participation in the study.

Will my taking part in this study be kept confidential?

Yes. We will follow ethical and legal practice guidelines and all information that is collected about you and your child will be handled in confidence. Neither you nor your child can be identified through the study results.

What will happen to the results of the research?

The results of this research will be published on the website, will be described to other doctors at scientific meetings and will be published in medical (peer-reviewed) journals. No individual children or families will be identifiable in any of this material. We sometimes use photos, without names, of children to illustrate our procedures. If we planned to do this, we would obtain your specific consent beforehand.

What happens to data collected in the study?

All data are stored securely (as set out in the Data Protection Act) and not released to any third party without your explicit permission. Unless you give permission for us to use the records for further research or teaching they are destroyed after the results of the study have been published.

Who can give me further information?

- You can read more about the study on our website:
 - http://[redacted]
- You can contact us by letter, telephone or email (see below) at any time and we will be very happy to answer any questions you may have.
- Finally **Bliss**, the premature baby charity, has a useful helpline and website to tell you about premature babies and their staff may be able to answer your questions.

Who is organising and funding the research?

The Study Director [redacted] is the lead researcher and main point of contact for the baby lab.

This study is funded by a national children’s medical research charity, **SPARKS**.

The Wellcome Trust supports the research into EEG analysis and our researchers are partly funded by the UCL Impact studentships and Great Ormond Street [redacted]

Who has reviewed the study?

The study was reviewed by a range of experts in the field during its development as we applied for funding. The study has been approved by the NW London Research Ethics Committee 2 (Reference 10/H0720/80) and is registered with the Research and Development Department of [redacted]

Useful Contacts

Chief Investigator: [redacted]

Contact details: [redacted]

[redacted] **Team:**

[redacted] D researcher

[redacted] PhD Researcher

Charlotte Sanderson, Trainee Clinical Psychologist

[redacted] Research Nurse

All contactable at the [redacted]

Additional contact details for [redacted]

queries): [redacted]

Bliss:

Parent Support Helpline

FREEPHONE 0500 618140

(Monday to Friday 10:00am - 5:00pm)

Website: www.bliss.org.uk

If you wish to complain about the conduct of this study please contact:

[redacted] **Patient Advice and Liaison Service**

By telephone: 020 3417 2042

By email: [redacted]

By letter: [redacted]

Appendix 8 – Autism Observation Schedule for Infants (AOSI)
(Bryson et al., 2008)

NOT REPRODUCED DUE TO COPYRIGHT RESTRICTIONS

Appendix 9 – Infant Behaviour Questionnaire Revised – Very Short Form (IBQ-R-VSF) (S.P. Putnam et al., 2014)

NOT REPRODUCED DUE TO COPYRIGHT RESTRICTIONS

(Can be requested free of charge at:
<http://www.bowdoin.edu/~sputnam/rothbart-temperament-questionnaires/request-forms/>)

Appendix 10 – Autism Diagnostic Observation Schedule II (Lord et al., 2012)

NOT REPRODUCED DUE TO COPYRIGHT RESTRICTIONS

Appendix 11 - Longitudinal Study Parental Consent Form

Patient Identification

Number for this trial: _____

Name of Researcher:

FORM FOR PARENTAL CONSENT

*Please initial
each box*

1. I confirm that I have read and understand the information sheet dated March 2015 (Version 4.0) for the above study. I have had the opportunity to consider the information, to ask questions and have had these answered satisfactorily.
2. I understand that the participation of my baby is voluntary and that I am free to withdraw at any time without giving any reason, without the medical care or legal rights of my baby being affected.
3. I understand that relevant sections of my baby's medical notes and data collected during the study may be looked at by individuals from regulatory authorities or from the NHS Trust, where it is relevant to my baby taking part in this research. I give permission for these individuals to have access to my baby's records.
4. I agree to the video recording of my child during the tests carried out in infancy for the purposes of scoring my child's response.
5. I agree that the video recordings and EEG tracings can be used for further research and teaching purposes; the material will always be used anonymously and my child will not be identifiable in the data used.
6. I agree to my GP being informed of our participation in the study.
7. I agree that my baby may take part in the above study.

Name of Child:		
Name of Parent:	Date	Signature
Name of Person taking consent:	Date	Signature

3 copies: one to be retained by parent, one placed in the clinical notes and one retained by the study office.

Version 4.0 March 2015

Appendix 12 – 30-34 Month Only Parental Consent Form (full term)

Participant study

ID Number:

Name of Researcher:

FORM FOR PARENTAL CONSENT (TERM BABY)

*Please initial
each box*

- 1 I confirm that I have read and understand the information sheet dated June 2015 (version 1.0) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
- 2 I understand that the participation of my child is voluntary and that I am free to withdraw at any time without giving any reason, without the medical care or legal rights of my child being affected.
- 3 I understand that relevant sections of my child's medical notes and data collected during the study may be looked at by individuals from regulatory authorities or from the NHS Trust, where it is relevant to my child taking part in this research. I give permission for these individuals to have access to my child's records.
- 4 I agree to the video recording of my child during the tests carried out for the purposes of scoring my child's response.
- 5 I agree that the video recordings and EEG tracings can be used for further research and teaching purposes; the material will always be used anonymously and my child will not be identifiable in the data used.
- 6 I give permission for the researchers involved in the Preterm Development Project to contact me in the future, should any further assessment stages go ahead.
- 7 I agree to my GP being informed of our participation in the study.
- 8 I agree that my child may take part in the above study.

Name of Child:		
Name of Parent:	Date	Signature
Name of Person taking consent:	Date	Signature

Appendix 13 – Parent Study Experience Feedback Form

Parents' feedback on research participation

We hope you and your baby enjoyed your visit to the Baby Lab today. Before you leave we would really appreciate you answering a few questions for us about your experience of taking part. Your answers are not part of the research and are purely to help improve the research experience for other parents and infants.

Date:

Baby's age (in months):

On a scale of 1 – 5, 1 being 'strongly disagree' and 5 being 'strongly agree', please answer the following questions about your visit to the lab today.

- | | | | | | |
|--|---|---|---|---|---|
| 1) My baby and I were made to feel welcome and comfortable | 1 | 2 | 3 | 4 | 5 |
| 2) The research tasks were explained fully to me and I felt able to ask questions if I needed to | 1 | 2 | 3 | 4 | 5 |
| 3) My baby remained content throughout the visit | 1 | 2 | 3 | 4 | 5 |
| 4) My visit was well-organised and ran smoothly | 1 | 2 | 3 | 4 | 5 |
| 5) I would recommend taking part to other friends and family with babies | 1 | 2 | 3 | 4 | 5 |

6) What, if anything, would have made your visit today more enjoyable?

.....
.....
.....
.....
.....
.....
.....

Appendix 14 – Demographic Questionnaire Section for Parent Education Level

1 What is your highest qualification from school or college?

	You	Your Partner
None of the below	<input type="checkbox"/> 1	<input type="checkbox"/> 1
Vocational qualification, NVQ, or CSE	<input type="checkbox"/> 2	<input type="checkbox"/> 2
O Level, GCSE, or Scottish Standards	<input type="checkbox"/> 3	<input type="checkbox"/> 3
BTEC, A Levels or Scottish Highers	<input type="checkbox"/> 4	<input type="checkbox"/> 4
Diploma or HND	<input type="checkbox"/> 5	<input type="checkbox"/> 5
Nursing qualification	<input type="checkbox"/> 6	<input type="checkbox"/> 6
University degree	<input type="checkbox"/> 7	<input type="checkbox"/> 7
Postgraduate University degree	<input type="checkbox"/> 8	<input type="checkbox"/> 8
Other qualification after A Level (please describe)	<input type="checkbox"/> 9	<input type="checkbox"/> 9